JOURNAL OF MATHEMATICS AND PHYSICS

Massachusetts Institute of Technology

Volume X

1930-1931

EDITORS

FREDERICK S. WOODS, Professor of Mathematics, Chairman
FREDERICK G. KEYES, Professor of Physico-Chemical Research
VANNEVAR BUSH, Professor of Electric Power Transmission
MANUEL S. VALLARTA, Associate Professor of Physics
CLARENCE L. E. MOORE, Professor of Mathematics

MANAGING EDITOR

PHILIP FRANKLIN, Associate Professor of Mathematics

QA 1
J 87
v. 10

OHIO STATE
UNIVERSITY

CONTENTS

PAGE

An Extension of Landau's Theorem in the Analytical Theory of Numbers 1
By SHIKAO IKEHARA

Some Formulas for the Strength of the Magnetic Field of a Cylindrical Coil 13
By ROBERT F. H. CHAO

The Geometry of Algebraic Pfaffians 19
By C. L. E. MOORE AND P. FRANKLIN

The Calculation of Resistances to Ground and of Capacitance 50
By HERBERT BRISTOL DWIGHT

On Tauberian Theorems of Hardy and Littlewood and a Note on Wintner's Paper 75
By SHIKAO IKEHARA

On Integral Functions with Real Negative Zeros 84
By SHIKAO IKEHARA

A Note on Infinite Products 92
By SHIKAO IKEHARA

Pfaffians in Parametric Form 95
By C. L. E. MOORE AND P. FRANKLIN

The Mapping Theorem of Koebe and the Problem of Plateau 106
By JESSE DOUGLAS

Stirling Expansions Derived by Means of Finite de La Vallee-Poussin Summation 131
By RAYMOND D. DOUGLASS

Geodesics of Pfaffians 157
By P. FRANKLIN AND C. L. E. MOORE

402250

PAGE

Synthesis of a Finite Two-terminal Network whose Driving-point Impedance is a Prescribed Function of Frequency 191
By Otto Brune

Letter to the Editor 237
By G. W. O. Howe

Dirac Equations in General Relativity
1. Four Dimensional Theory 239
By J. A. Schouten

Dirac Equations in General Relativity
2. Five Dimensional Theory 272
By J. A. Schouten

A New Deduction of the Gaussian Distribution . . . 284
By Norbert Wiener

Algebraic Matric Equations 289
By Philip Franklin

The Problem of Plateau for Two Contours 315
By Jesse Douglas

AN EXTENSION OF LANDAU'S THEOREM IN THE ANALYTICAL THEORY OF NUMBERS

By Shikao Ikehara

I. Introduction

Abel's theorem states that if $\sum a_n$ is convergent, then $\lim \sum a_n x^n$ exists as $x \to 1$ by real values; and is equal to $\sum a_n$. The converse theorem, however, that the existence of $\lim_{x \to 1-0} \sum a_n x^n$ implies the convergence of $\sum a_n$ is not unconditionally true. Tauber [1] showed that if $\lim_{x \to 1-0} \sum a_n x^n = A$ and $\lim n a_n = 0$, it follows that $\sum a_n = A$. Hardy and Littlewood [2] have shown that $\lim_{x \to 1-0} \sum a_n x^n = A$ and the auxiliary condition that $n a_n > -K$ form a sufficient hypothesis for this condition: $\sum a_n = A$. The term "Tauberian theorems" is applied to theorems of this type, namely those which involve some hypothesis on the rapidity of growth of the terms of a series, and under this assumption proceed in the ordinary sense from the given summability of the series by some particular method to its convergence.

The methods of Hardy and Littlewood are difficult, involving as they do repeated differentiations and they are furthermore but little systematized, so that each particular theorem involves a considerable amount of manipulation peculiar to itself. Robert Schmidt [3] attacked the problem by means of theory of moments and gave a general theory of Tauberian theorems adequate to deal with the majority of known cases.

Introducing the method of Fourier analysis, Professor Wiener [4] brought about not only a systematization, but a simplification of a large portion of the field in the theory of Tauberian theorems. He puts Tauberian theorems in the form: if a function has a mean value of a certain type for large values of its arguments, and if

PHYSICS LIBRARY

the rate of growth of the function is subject to certain restrictions, then the function has a limiting value at infinity. Closely related to theorems of this type are theorems which read as follows: if a function has a mean of a certain given type at infinity, and its growth is restricted in a certain given manner, then it has a mean of a certain other given type at infinity.

By means of theorems of this type, he applied his method to prove the O_L Tauberian theorems for Abel, Lambert, and Borel summation, and the Cesàro summability of the Fourier integral.[5] Moreover, his proof of the theorem on Lambert summation practically proves the classical theorem of the distribution of primes, and is the first method of proof that does not assume the validity of the theorem first. S. B. Littauer [6] developed other theorems of this character which may be applied to questions of the summability of Fourier series and integral.

The author wishes to express deep gratitude to Professor Norbert Wiener for his encouragement and guidance during the progress of this paper.

II. Wiener's Tauberian Theorems

For our purpose we state Professor Wiener's Tauberian theorems as follows: *Let* $\{M_k(x)\}$ *be a class of real functions which are* $O\left(\frac{1}{x(\log x)^2}\right)$ *both at* 0 *and at* ∞, *and let* $M_2(x)$ *be a real function which is also* $O\left(\frac{1}{x(\log x)^2}\right)$ *at* 0, *and at* ∞, *and which are both bounded over every interval* $(\epsilon, 1/\epsilon)$. *Let* $\{M_k(x)\}$ *be non-negative and measurable, and let any real number* A *be given, then there is a* k *such that*

$$\int_0^\infty M_k(x)x^{iu}dx$$

is not zero for

$$-A < u < A.$$

Let $M_2(x)$ *be continuous except for a finite number of finite jumps.*

Let $\varphi(x)$ be a non-negative measurable function such that

(a) $$\lim_{y \to 0} \int_0^\infty \phi\left(\frac{x}{y}\right) M_k(x)dx = C \int_0^\infty N_k(x)dx;$$

(b) $$\int_0^\infty \phi\left(\frac{x}{y}\right) M_k(x)dx \text{ is bounded.} \quad (0 < y < \infty)$$

Then

(c) $$\lim_{y \to 0} \int_0^\infty \phi\left(\frac{x}{y}\right) M_2(x)dx = C \int_0^\infty M_2(x)dx.$$

Following Professor Wiener's method of proof we shall show that our extension leaves his theorem still valid.

Introducing transformations

$$x = e^\xi; \qquad y = e^\eta; \qquad xM_k(x) = N_k(\xi);$$
$$xM_2(x) = N_2(\xi); \qquad \phi(x) = f(\xi);$$

the theorem can be put in the form: *Let $\{N_k(\xi)\}$ be a class of bounded functions which are $O(1/\xi^2)$ at $\pm \infty$, and let $N_2(\xi)$ be a bounded function which is also $O(1/\xi^2)$ at $\pm \infty$. Let $N_k(\xi)$ be measurable and non-negative, and let any real number A be given, then there is a k such that*

$$\int_{-\infty}^\infty N_k(\xi) e^{iu\xi} d\xi$$

is not zero for

$$-A < u < A.$$

Let $N_2(\xi)$ be continuous, except for a finite number of finite jumps. Let $f(\xi)$ be a non-negative measurable function such that

(a) $$\lim_{\eta \to -\infty} \int_{-\infty}^\infty f(\xi - \eta) N_k(\xi) d\xi = C \int_{-\infty}^\infty N_k(\xi) d\xi;$$

(b) $$\int_{-\infty}^\infty f(\xi - \eta) N_k(\xi) d\xi \text{ is bounded.} \quad [-\infty < \eta < \infty]$$

Then

(c) $$\lim_{\eta \to -\infty} \int_{-\infty}^\infty f(\xi - \eta) N_2(\xi) d\xi = C \int_{-\infty}^\infty N_2(\xi) d\xi.$$

Let $R(\xi)$ be an absolutely integrable function, then we shall show that

$$N(\xi) = \int_{-\infty}^{\infty} N_k(\xi - \zeta)R(\zeta)d\zeta$$

satisfies (c).

Since R is absolutely integrable and both f and N_k are non-negative, we can invert the order of integration, i.e.,

$$\int_{-\infty}^{\infty} f(\xi - \eta)N(\xi)d\xi = \int_{-\infty}^{\infty} f(\xi - \eta)d\xi \int_{-\infty}^{\infty} N_k(\xi - \zeta)R(\zeta)d\zeta$$

$$= \int_{-\infty}^{\infty} R(\zeta)d\zeta \int_{-\infty}^{\infty} f(\xi - \eta + \zeta)N_k(\xi)d\xi.$$

Write

$$\int_{-\infty}^{\infty} f(\xi - \eta)N_k(\xi)d\xi - C\int_{-\infty}^{\infty} N_k(\xi)d\xi = \psi(\eta);$$

the function is then a bounded function approaching 0 as η tends toward $-\infty$.

We have

$$\left|\int_{-\infty}^{\infty} R(\zeta)\psi(\eta - \zeta)d\zeta\right|$$

$$\leq \left|\int_{-\infty}^{B} R(\zeta)\psi(\eta - \zeta)d\zeta\right| + \left|\int_{B}^{\infty} R(\zeta)\psi(\eta - \zeta)d\zeta\right|$$

$$\leq \max|\psi(\eta)| \int_{-\infty}^{B} |R(\zeta)|d\zeta$$

$$+ \max_{\zeta \geq B}|\psi(\eta - \zeta)| \int_{-\infty}^{\infty} |R(\zeta)|d\zeta.$$

It follows that

$$\overline{\lim_{\eta \to -\infty}} \left|\int_{-\infty}^{\infty} R(\zeta)\psi(\eta - \zeta)d\zeta\right| \leq \max|\psi(\eta)| \int_{-\infty}^{B} |R(\zeta)|d\zeta,$$

and as B is arbitrary, that

$$\overline{\lim_{\eta \to -\infty}} \left|\int_{-\infty}^{\infty} R(\zeta)\psi(\eta - \zeta)d\zeta\right| = 0,$$

or as we may write it, that

$$\overline{\lim_{\eta \to -\infty}} \int_{-\infty}^{\infty} f(\xi - \eta) N(\xi) d\xi = C \int_{-\infty}^{\infty} R(\zeta) d\zeta \int_{-\infty}^{\infty} N_k(\xi) d\xi$$
$$= C \int_{-\infty}^{\infty} N(\xi) d\xi.$$

The boundedness of N follows from that of N_k and the absolute integrability of R. Thus any function $N(\xi)$ satisfies (c).

Let us put

$$h_k(u) = \int_{-\infty}^{\infty} N_k(\xi) e^{iu\xi} d\xi$$

which we assume not zero for

$$-A \leqslant u < A.$$

By the Plancherel theory of the Fourier transform,

$$g_k(u) = \underset{B \to \infty}{\text{l.i.m.}} \int_{-B}^{B} i\xi N_k(\xi) e^{iu\xi} d\xi$$

exists as a measurable function the square of whose modulus is Lebesgue integrable from $-\infty$ to ∞.[7] We have

$$\int_{\alpha}^{\beta} g_k(u) du = \int_{-\infty}^{\infty} i\xi N_k(\eta) \left[\int_{\alpha}^{\beta} e^{iu\xi} du \right] d\xi = h_k(\beta) - h_k(\alpha).$$

The function $h_k(u)$ is thus a non-vanishing function in the finite region of u.

Let $N(\xi)$ be a bounded function which is $O(1/\xi^2)$ at $\pm \infty$. Let

$$s(u) = \int_{-\infty}^{\infty} N(\xi) e^{iu\xi} d\xi,$$

and let

$$t(u) = \underset{B \to \infty}{\text{l.i.m.}} \int_{-B}^{B} i\xi N(\xi) e^{iu\xi} d\xi.$$

As above, we have

$$\int_{\alpha}^{\beta} t(u) du = s(\beta) - s(\alpha).$$

Let us make the further restriction that

$$t(u) = 0 \qquad [|u| \geq A].$$

The proof is as follows: consider

$$\omega(u) = \frac{h_k(u)t(u) - g_k(u)s(u)}{[h_k(u)]^2}.$$

It is seen that $s(u)$ as well as $t(u)$ must vanish for $|u| \geq A$ as it is constant outside this range and is quadratically summable. Over the range $(-A, A)$, $h_k(u)$ and $s(u)$ are bounded continuous functions, but $h_k(u)$ vanishes outside the range $(-A, A)$.

Let $h_k(u)$ have the Fourier series $\sum_{-\infty}^{\infty} a_n e^{inuk}$. Then $g_k(u)$ will have the Fourier series $\sum in\, a_n e^{inu}$. Similarly, let $s(u)$ have the Fourier series $\sum b_n e^{inu}$, then $t(u)$ will have the Fourier series $\sum in\, b_n e^{inu}$.

Let

$$h_k^N(u) = \sum_{-N}^{N} a_n e^{inkn}\left(1 - \frac{|n|}{N}\right);$$

$$g_k^N(u) = \sum_{-N}^{N} ink a_n e^{inku}\left(1 - \frac{|u|}{N}\right);$$

$$s^N(u) = \sum_{-\infty}^{N} b_n e^{inu}\left(1 - \frac{|u|}{N}\right);$$

$$t^N(u) = \sum_{-N}^{N} in b_n e^{inu}\left(1 - \frac{|u|}{N}\right).$$

Then we can write

$$\int_{-\pi}^{\pi}\left|\omega(u) - \frac{h_k^N(u)t^N(u) - g_k^N(u)s^N(u)}{[h_k^N(u)]^2}\right|^2 du$$

$$= \int_{-\pi}^{\pi}\left|\frac{h_k(u)t(u)[h_k^N(u)]^2 - g_k(u)s(u)[h_k^N(u)]^2 - h_k^N(u)t^N(u)[h_k(u)]^2 + g_k^N(u)s^N(u)[h_k(u)]^2}{[h_k(u)]^2[h_k^N(u)]^2}\right| du.$$

The functions $h_k^N(u)$, $s^N(u)$ and $1/h_k^N(u)$ are all uniformly bounded in u and N and converge uniformly to the corre-

sponding functions without the N. Then it can be shown that

$$\text{l.i.m.}_{N\to\infty} \frac{h_k^N(u)t^N(u) - g^N(u)s^N(u)}{[h_k^N(u)]^2} = \omega(u).$$

Similarly,

$$\lim_{N\to\infty} \frac{s^N(u)}{h_k^N(u)} = \frac{s(u)}{h_k(u)}$$

uniformly in u. Since

$$\int_\alpha^\beta \frac{h_k^N(u)t^N(u) - g^N(u)s^N(u)}{[h_k^N(u)]^2}\,dn = \frac{s^N(\beta)}{h_k^N(\beta)} - \frac{s^N(\alpha)}{h_k^N(\alpha)},$$

we have

$$\int_\alpha^\beta \omega(u)du = \frac{s(\beta)}{h_k(\beta)} - \frac{s(\alpha)}{h_k(\alpha)}.$$

Write

$$R(\xi) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{s(u)}{h_k(u)} e^{-iu\xi}du,$$

then

$$N(\xi) = \int_{-\infty}^{\infty} N_k(\xi - \zeta)R(\zeta)d\zeta.$$

Also

$$R(\xi) = \frac{1}{2\pi i\xi}\int_{-\infty}^{\infty} \omega(u)e^{-iu\xi}du.$$

$\int_{-\infty}^{\infty} \omega(u)e^{-iu\xi}du$ is quadratically summable; hence, by means of the Schwartz inequality we can show that

$$\int_B^\infty |R(\xi)|\,d\xi \qquad \text{and} \qquad \int_{-\infty}^{-B} |R(\xi)|\,d\xi$$

are bounded. From the above $R(\xi)$ is absolutely integrable from $-\infty$ to ∞. Thus $N_k(\xi)$ satisfies (c).

III. On Landau's Theorem

We propose to prove the theorem:

Let $\varphi(x + iy) = \sum_1 a_n n^{-x-iy}$ *converge absolutely when* $x \to 1$. *Let the analytic function* φ *be continued on to the line* $x = 1$, *and*

let it be analytic there, except for a pole of order 1 *with residue* A *at* $x + iy = 1$, *then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{1}^{N} a_k = A.$$

Landau [8] proved this theorem with the aid of the additional condition that:

$$\varphi(x + iy) = O(|y|^k).$$

We shall proceed to prove that the other condition is not necessary to the proof of the above theorem.

Let us first consider a generalized theorem of Hardy and Littlewood,[9] which reads:

Suppose that (i) *the series* $\sum a_n\lambda_n^{-s}$ *is absolutely convergent for* $\sigma > \sigma_0 > 0$; (ii) *the function* $F(s)$ *defined by the series is regular for* $\sigma > c$, *where* $0 < c \leq \sigma_0$, *and continuous for* $\sigma \geq c$; (iii) $F(s) = O(e^{|t|})$, *where* $C < \frac{1}{2}\pi$, *uniformly for* $\sigma \geq c$; (iv) $\lambda_n/\lambda_{n-1} \to 1$; (v) a_n *is real, and satisfies one or other of the inequalities.* $a_n > -K\lambda_n^{c-1}(\lambda_n - \lambda_{n-1})$, $a_n < K\lambda_n^{c-1}(\lambda_n - \lambda_{n-1})$, *or is complex and of the form* $O\{\lambda_n^{c-1}(\lambda_n - \lambda_{n-1})\}$; *and* (vi) $F(s)$ *has a simple pole at the point* $s = c$, *and that the residue at the pole is* g. *Then*

$$A_n = a_1 + a_2 + a_3 + \cdots + a_n \sim \frac{g\lambda_n^c n}{c}.$$

If we suppose that in the above theorem $a_n \geq 0$, that $F(s)$ is regular on the line $\sigma = c$, and that the equation $F(s) = O(e^{c(t)})$ is replaced by $F(s) = O(|t|^k)$, the theorem reduces to that of Landau.

We shall show that the theorem of Hardy and Littlewood can be proved without use of the conditions (iii), and (v).

Consider that:

$$\varphi(x) = \sum_{1} a_n n^{-x}.$$

Integrating this function from x to b(x and $b > 1$), we have

$$\int_x^b \varphi(x)dx = \frac{\sum a_n n^{-x}}{(\log n)} - \frac{\sum a_n n^{-b}}{(\log n)}, \tag{1}$$

In this integration we changed a pole to a logarithmic singularity. If we integrate (1) again from x to b (x and $b > 1$), we obtain:

$$\frac{\sum a_n n^{-s}}{(\log n)^2} - \frac{\sum a_n n^{-b}}{(\log n)} - \frac{\sum a_n n^{-b}}{(\log n)}(b - x). \tag{2}$$

It will be noticed that the second integration transformed logarithmic singularity to a finite quantity; in other words, the whole expression of (2) is bounded. The second and third term in (2) converge for a sufficiently large value of b. The first term is thus a bounded monotone non-increasing function. Hence it should reach a definite limit as $x \to 1$; i.e.,

$$\lim_{s \to 1} \frac{\sum a_n n^{-s}}{(\log n)2} = C. \tag{3}$$

Next, we evaluate

$$\frac{1}{2\pi}\int_{-B}^{B}\left(1 - \frac{|y|}{B}\right) n^{-s-iy}e^{-iuy}dy$$

$$= \frac{n^{-s}}{\pi}\int_{0}^{B}\left(1 - \frac{y}{B}\right)\cos(u + \log n)y\,dy$$

$$= \frac{n^{-s}}{\pi B(n + \log n)^2}[1 - \cos(u + \log n)B]. \tag{4}$$

Writing

$$\varphi(x + iy) = \sum a_n n^{-s-iy}.$$

We have from (4)

$$\frac{1}{2\pi}\int_{-B}^{B}\left(1 - \frac{|y|}{B}\right)\varphi(x + iy)e^{-iuy}dy$$

$$= \sum \frac{a_n n^{-s}(1 - \cos B(u + \log n))}{\pi B(u + \log n)^2} \tag{5}$$

when $x \to 1$. Since this series is less term by term than the convergent series (3), this is certainly convergent.

The expression (3) can be put in the form:

$$\frac{1}{2\pi}\int_0^B \left(1-\frac{y}{B}\right)\{[\varphi(x+iy)+\varphi(x-iy)]\cos uy - i[\varphi(x+iy)-\varphi(x-iy)]\sin uy\}dy.$$

Placing $x = 1$, we get finally

$$\frac{1}{2\pi}\int_0^B \left(1-\frac{y}{B}\right)\{[\varphi(1+iy)+\varphi(1-iy)]\cos uy - i[\varphi(1+iy)-\varphi(1-iy)]\sin uy\}dy = \sum \frac{a_n[1-\cos B(u+\log u)]}{a\pi B(n+\log n)^2}. \quad (6)$$

Let

$$\varphi(1+iy) = \frac{A}{iy} + g(iy),$$

where $g(iy)$ is an analytic function. Then

$$\varphi(1+iy)+\varphi(1-iy) = g(iy)+g(-iy) = G(iy).$$

Hence:

$$\begin{aligned}\int_0^B [\varphi(1+iy)+\varphi(1-iy)]\cos uydy &= \int_0^B G(iy)\cos uydy \\ &= \frac{1}{u}\int_0^B G(iy)d(\sin uy) \\ &= \frac{1}{u}\int_0^B G(iy)\sin uydy - \frac{1}{u}\int_0^B \sin uyG'(iy)dy \\ &= o\left(\frac{1}{u}\right).\end{aligned}$$

Similarly:

$$\begin{aligned}\varphi(1+iy)-\varphi(1-iy) &= \frac{A}{iy}+g(iy)+\frac{A}{iy}-g(-iy) \\ &= \frac{2A}{iy}+H(iy)\end{aligned}$$

and therefore

$$
\begin{aligned}
-i\int_0^B [\varphi(1+iy)-\varphi(1-iy)]\sin uydy &= -i\int_0^B\left[\frac{2A}{iy}+H(iy)\right]\sin uydy \\
&= -2A\int_0^B \frac{\sin uy}{y}dy + i\int_0^B H(iy)\sin uydy \\
&= -2A\int_0^B \frac{\sin uy}{uy}d(uy) + o\left(\frac{1}{u}\right).
\end{aligned}
$$

Since

$$\lim_{u\to\infty}\frac{-A}{\pi}\int_0^\infty \frac{\sin uy}{uy}d(uy) = \frac{A}{2};$$

we have from (6)

$$\frac{\sum a_n[1-\cos B(u+\log n)]}{n\pi B(u+\log n)^2} = \lim_{u\to-\infty}\frac{-A}{2\pi}\int_0^B \frac{\sin uy}{uy}duy = \frac{A}{2}. \quad (7)$$

If we place $u = \log x$ in (7), we have:

$$\lim_{u\to\infty}\sum_{n=1}\frac{2a_n}{n\pi}\left(\frac{1-\cos B\log xn}{B(\log xn)^2}\right) = A, \quad (8)$$

which is true for every B. It may be mentioned here that the only restriction placed on a_n is that it be positive.

To sum up our argument: We have shown in (5) that the class of functions we are dealing with is bounded. Since the weighting factor is $(1-|y|/B)$, the Fourier transform of $\varphi(x+iy)$ in (4) is zero from a certain point on. But we may extend the range as far as required by increasing B. Moreover,

$$\left(\frac{1-\cos B\log xn}{B(\log xn)^2}\right)$$

in (7) is a bounded measurable function, which is $O(1/(xn)^2)$ at 0 and ∞.

Hence we have

$$\lim_{N\to\infty}\frac{1}{N}\sum_1^N a_k = A.$$

Thus Wiener's condition is satisfied: and Landau's theorem is proved without making use of the additional condition. Moreover, Hardy and Littlewood's theorem is proven without use of the conditions (iii) and (v).

As an immediate consequence of our theorem, we can prove the prime number theorem. Suppose that $\Lambda(n)$ is zero unless n is a power of a prime p, in which case $\Lambda(n)$ is equal to $\log p$.

Since

$$\sum_{n=1}^{\infty} \Lambda(n) n^{-s-iy} = -\frac{\zeta'(x+iy)}{\zeta(x+iy)}$$

and since this function has a pole of order 1 with residue 1 at $x + iy = 1$,

$$\frac{1}{n} \sum_{1}^{n} \Lambda(k) \to 1.$$

IV. Bibliography

1. A. Tauber, Ein Satz aus der Theorie der unendlichen Reichen, *Monatsch. f. Math. u. Phys.* 8 (1897) pp. 273–277.
2. Hardy and Littlewood, Tauberian Theorems concerning Power Series and Dirichlet's Series whose Coefficients are Positive. *Proc. Lond. Math. Soc.*, Series 2, 13 (1913).
3. Schmidt, R., Über divergente Folgen und lineare Mittlebildungen. *Math. Ztschr.* 22 (1925), pp. 89–152.
4. N. Wiener, A New Method in Tauberein Theorems, *Jour. of Math. and Phys., Mass. Inst. of Tech.*, 7 (1928), pp. 161–184.
5. N. Wiener, A Type of Tauberian Theorem Applying to Fourier Series, *Proc. Lond. Math. Soc.*, Series 2, 30 (1929), pp. 1–8.
6. S. B. Littauer, A New Tauberian Theorem with Application to the Summability of Fourier Series and Integrals, *Jour. of Math. and Phys., Mass. Inst. of Tech.*, 8 (1929), pp. 216–234.
7. M. Plancherel, contribution à l'étude de la représentation d'une fonction arbitraire *par* des intégrales définies. *Rendiconti di Palermo*, 30 (1910), pp. 289–335.
8. E. Landau, Handbuch der Lehre von der Verteilung der Primzahlen. p. 874.
   For a thorough historical account of this subject, see H. Bohr and H. Cramér, *Enc. Math. Wiss.*, II C 8.
9. Hardy and Littlewood, Proofs of the Prime—Number Theorem and Similar Theorems, *Quart. Jour.* 46 (1925), pp. 215–219.
   Contributions to the theory of the Riemann Zetafunction and the Theory of the Distribution of Primes, *Acta Math.* 41 (1918), pp. 119–196.

SOME FORMULAS FOR THE STRENGTH OF THE MAGNETIC FIELD OF A CYLINDRICAL COIL

By Robert F. H. Chao

The strength of the magnetic field near a circular coil or solenoid, in a non-magnetic medium, is a useful quantity to compute in several classes of electrical work. In work with laboratory instruments the strength of field near a coil is often required to be known. In designing magnetic shields for circular coils, the field in a non-magnetic medium is required. In computing the effect of current-limiting reactors on structural steel members near them, the strength of the magnetic field is of importance.

In this paper, a number of formulas for the strength of the magnetic field are derived, which cover a large part of the field of a coil, and comparisons are made with published formulas.

The simplest formulas are those for the strength of field of a single, circular filament carrying current, and these will first be given.

The magnetic potential near a circular filament carrying a current of I amperes is [1]

$$\frac{\omega I}{10} = \frac{2\pi I}{10} \sin^2 \alpha \sum_{n=1}^{\infty} \frac{1}{n+1} P_n'(\cos \alpha) P_n(\mu) \left(\frac{\rho}{r}\right)^{n+1}, \qquad (1)$$

where the dimensions are in cm. and are as indicated in Figure 1, where

$$\mu = \frac{x}{r} = \frac{x}{\sqrt{x^2 + y^2}}$$

and where P_n and P_n' denote zonal harmonics and their derivatives. See Appendices I and II. The letter n denotes integers.

[1] Absolute Measurements in Electricity and Magnetism, by Andrew Gray, ed. of 1893, Vol. II, Part I, page 47, eq. (82).

If the circular filament have its center at the origin of coördinates, as in Figure 2, the terms involving even values of n vanish and eq. (1) becomes

$$\frac{\omega I}{10} = \frac{\pi I}{10}\left[\frac{a^2}{r^2}P_1(\mu) - \frac{3}{4}\frac{a^4}{r^4}P_3(\mu) + \frac{3.5}{4.6}\frac{a^6}{r^6}P_5(\mu) - \cdots\right], \quad (2)$$

where $r^2 = x^2 + y^2$. By expressing the zonal harmonics as

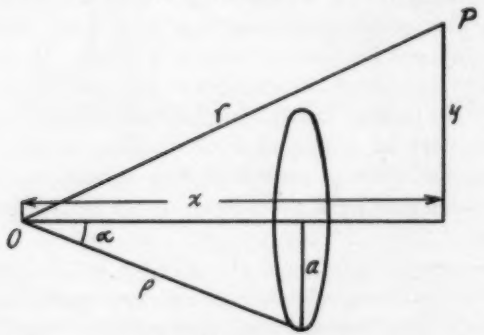


FIG. 1.—Circular Filament.

expressions in x and y and differentiating with respect to x, the magnetic field component parallel to the axis of the circle is obtained. It is

$$-\frac{I}{10}\frac{\partial \omega}{\partial x} = \frac{2\pi I}{10r}\left[\frac{a^2}{r^2}P_2(\mu) - \frac{3}{2}\frac{a^4}{r^4}P_4(\mu) + \frac{3.5}{2.4}\frac{a^6}{r^6}P_6(\mu) - \cdots\right]. \quad (3)$$

Tables can be used for the numerical values of the zonal harmonics.[2]

By differentiating with respect to y, the radial component of the magnetic field is obtained. It is:

$$-\frac{I}{10}\frac{\partial \omega}{\partial y} = \frac{\pi I y}{10r^2}\left[\frac{a^2}{r^2}P_2'(\mu) - \frac{3}{4}\frac{a^4}{r^4}P_4'(\mu) + \frac{3.5}{4.6}\frac{a^6}{r^6}P_6'(\mu) - \cdots\right]. \quad (4)$$

Equations (3) and (4) are suitable for points at a distance from

[2] See "Fourier's Series and Spherical Harmonics," by W. E. Byerly, First Edition, page 280.

the center of the circle so that a/r is less than approximately 1. For points near the center of the circle, or near the axis, Gray's formulas (9) and (10) can be used.[3]

Equation (3) can be checked by means of a formula for the mutual inductance of two circular filaments with parallel axes.[4]

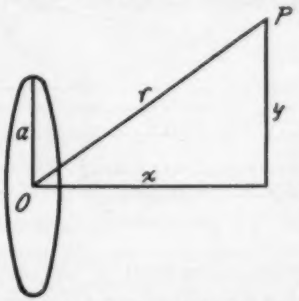


FIG. 2.—Circular Filament with Center at Origin.

By dividing by πa^2 and then putting the smaller radius equal to zero so that $K_n = 1$, eq. $(10A)$ becomes the same as eq. (3) of this paper.

Example: Find the radial component of the magnetic field near a single turn of fine wire carrying 10 amperes (See Fig. 2). $a =$ 2 cm., $x = 8$ cm. and $y = 1$ cm. By eq. (4),

$$-\frac{I}{10}\frac{\partial \omega}{\partial y} = 0.04833(0.1832 - 0.0275 + 0.0028)$$
$$= 0.00766.$$

By eq. (10), Reference 3,

$$-\frac{I}{10}\frac{\partial \omega}{\partial y} = 0.00791(1 - 0.0330 + 0.0007)$$
$$= 0.00766.$$

[3] "Absolute Measurements in Electricity and Magnetism," by Andrew Gray, ed. of 1893, Volume II, Part I, page 248 and ed. of 1921, page 212, equations (9) and (10). Change $1680x^4$ to $1680a^2x^4$. A simplification can be made by taking a factor 7 out of the last bracket of each formula.

[4] S. Butterworth, *Phil. Mag.*, Volume 31, 1916, page 443 and Scientific Paper of the Bureau of Standards No. 320, eq. 10 A.

Expressions for the magnetic field of a thick solenoid can be obtained by integrating equations (3) and (4) over the cross section of the solenoid, which is assumed to have uniform current

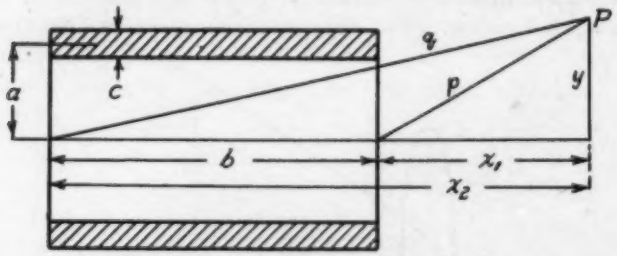


FIG. 3.—Cylindrical Coil or Solenoid.

density. The radical component of the magnetic field of a cylindrical coil whose dimensions are indicated in Fig. 3, is

$$R = \frac{\pi NIy}{10ab}\left[\left\{ \frac{a^3}{p^3} P_1'(\mu_1) - \frac{a^3}{q^3} P_1'(\mu_2) \right\}\left(1 + \frac{c^2}{12a^2} \right) - \frac{1}{4}\left\{ \frac{a^5}{p^5} P_3'(\mu_1) - \frac{a^5}{q^5} P_3'(\mu_2) \right\}\left(1 + \frac{c^2}{2a^2} + \frac{c^4}{80a^4} \right) + \frac{1.3}{4.6}\left\{ \frac{a^7}{p^7} P_5'(\mu_1) - \frac{a^7}{q^7} P_5'(\mu_2) \right\} \times \left(1 + \frac{5}{4}\frac{c^2}{a^2} + \frac{3}{16}\frac{c^4}{a^4} + \frac{c^6}{448a^6} \right) - \cdots \right] \quad (5)$$

where

$$p^2 = x_1^2 + y^2,$$
$$q^2 = x_2^2 + y^2,$$

$$\mu_1 = \frac{x_1}{p}$$

and

$$\mu_2 = \frac{x_2}{q}.$$

The corresponding expression for the axial component of the field of a cylindrical coil has been published.[5] It is as follows,

[5] "The Force between Unequal Reactance Coils with Parallel Axes," by H. B. Dwight and R. W. Pursell, *General Electric Review*, July, 1930, page 401.

the notation being changed to that of this paper:

$$F = \frac{\pi NI}{10b}\left[\left\{\frac{a^2}{p^2}P_1(\mu_1) - \frac{a^2}{q^2}P_1(\mu_2)\right\}\left(1 + \frac{c^2}{12a^2}\right)\right.$$
$$- \frac{3}{4}\left\{\frac{a^4}{p^4}P_3(\mu_1) - \frac{a^4}{q^4}P_3(\mu_2)\right\}\left(1 + \frac{c^2}{2a^2} + \frac{c^4}{80a^4}\right)$$
$$+ \frac{3.5}{4.6}\left\{\frac{a^6}{p^6}P_5(\mu_1) - \frac{a^6}{q^6}P_5(\mu_2)\right\}$$
$$\left.\times\left(1 + \frac{5}{4}\frac{c^2}{a^2} + \frac{3}{16}\frac{c^4}{a^4} + \frac{c^6}{448a^6}\right) - \cdots\right]. \quad (6)$$

If, in Fig. 3, the point P lie between the planes of the ends of the coil, x_1 and μ_1 are negative.

For the zonal harmonics, see the appendices, or values may be taken from tables (See Reference 2). The last bracket in each term is

$$\left\{\left(1 + \frac{c}{2a}\right)^n - \left(1 - \frac{c}{2a}\right)^n\right\}\frac{a}{cn}, \quad (7)$$

where n is an odd number.

Expressions (5) and (6) are suitable for points in any direction, at a distance from the center of the coil, so that a/p and a/q are less than approximately 1. For computing the axial component of the field near the center of the coil or near the axis, equations (28) and (31) of Gray can be used.[6]

The work described herein was done in connection with a thesis in the Department of Electrical Engineering at Massachusetts Institute of Technology under the direction of Dr. H. B. Dwight.

[6] "Absolute Measurements in Elec. and Mag.," by A. Gray, ed. of 1921, pages 222 and 225. Change $5a$ to $5a^2$. Note that

$$r_1^2 = a^2 + \left(x + \frac{b}{2}\right)^2 \quad \text{and} \quad r_2^2 = a^2 + \left(x - \frac{b}{2}\right)^2.$$

Use only the ed. of 1921 for eq. 31.

Appendix I—Zonal Harmonic Functions.

$$P_0(\mu) = 1,$$

$$P_1(\mu) = \mu,$$

$$P_2(\mu) = \frac{1}{2}(3\mu^2 - 1),$$

$$P_3(\mu) = \frac{1}{2}(5\mu^3 - 3\mu),$$

$$P_4(\mu) = \frac{1}{2.4}(5.7\mu^4 - 2.3.5\mu^2 + 1.3),$$

$$P_5(\mu) = \frac{1}{2.4}(7.9\mu^5 - 2.5.7\mu^3 + 3.5\mu),$$

$$P_6(\mu) = \frac{1}{2.4.6}(7.9.11\mu^6 - 3.5.7.9\mu^4 + 3.3.5.7\mu^2 - 1.3.5),$$

$$P_7(\mu) = \frac{1}{2.4.6}(9.11.13\mu^7 - 3.7.9.11\mu^5 + 3.5.7.9\mu^3 - 3.5.7\mu).$$

. . . . . .

Note that the brackets contain binomial coefficients as well as other factors.

Appendix II—First Derivatives.

$$P_n'(\mu) = \frac{d}{d\mu} P_n(\mu).$$

$$P_0'(\mu) = 0,$$

$$P_1'(\mu) = 1,$$

$$P_2'(\mu) = 3\mu,$$

$$P_3'(\mu) = \frac{1}{2}(3.5\mu^2 - 1.3),$$

$$P_4'(\mu) = \frac{1}{2}(5.7\mu^3 - 3.5\mu),$$

$$P_5'(\mu) = \frac{1}{2.4}(5.7.9\mu^4 - 2.3.5.7\mu^2 + 1.3.5),$$

$$P_6'(\mu) = \frac{1}{2.4}(7.9.11\mu^5 - 2.5.7.9\mu^3 + 3.5.7\mu),$$

$$P_7'(\mu) = \frac{1}{2.4.6}(7.9.11.13\mu^6 - 3.5.7.9.11\mu^4 + 3.3.5.7.9\mu^2 - 1.3.5.7),$$

. . . . . .

THE GEOMETRY OF ALGEBRAIC PFAFFIANS[1]

By C. L. E. Moore and P. Franklin

1. Introduction. If a Pfaffian expression in n variables be equated to zero, and the variables interpreted as co-ordinates in n-space, there is determined a field of "planar" elements analogous to the tangent "planes" ($n-1$ spaces) to a family of "surfaces" ($n-1$ dimensional manifolds). This geometry has already been studied to some extent by one of the authors.[2] The object of the present paper is to provide a basis for further generalization in this direction, by studying in detail certain simple, concrete examples. Accordingly we confine our attention to Pfaffians whose coefficients are polynomials in the variables, of at most the second degree, and place our emphasis on the three-dimensional case.

In the three-dimensional case, when the coefficients are of the first degree, there is a congruence of straight lines among the integral curves. A classification of these congruences is made, as well as a description of a set of normal forms to which all cases may be reduced by affine transformations. The behavior under projective transformations is also discussed, and the appropriate set of homogeneous variables for Pfaffians defined.

2. The two-dimensional affine case. Since the vanishing of a Pfaffian in two variables,

$$A(x, y)\,dx + B(x, y)\,dy = 0, \tag{2.1}$$

is an ordinary differential equation, it possesses an integrating factor, and the linear elements are simply those of a family of curves. In spite of the fact that here all cases are integrable,

[1] Presented to the American Mathematical Society, September, 1930.

[2] Cf. C. L. E. Moore, this *Journal*, vol. IX, pp. 320–332, J. A. Schouten, *Math. Zeitschrift*, vol. 30, 1929, pp. 149–172, and the references there given.

we shall find it suggestive for the later developments to discuss the equation (2.1).

If A and B are constants, the integral curves are a family of parallel lines, and the only singular points are at infinity.

If A and B are each polynomials of the first degree in x and y, it is well known that the equations of the integral curves may be found by elementary methods. We may obtain more insight into the nature of the integral curves, however, by reducing equation (2.1) to certain normal forms. We do this by suitably chosen affine transformations of the x, y plane.

Let us first note that we may rewrite equation (2.1) in the form

$$dQ + C(xdy - ydx) = 0, \tag{2.2}$$

where dQ is an exact differential. If the curves

$$Q = k \tag{2.3}$$

are central conics, they have the same principal axes. If they are parabolas, they have the same axis. In both cases, as well as the degenerate cases when Q is not a true conic, a rotation of our axes will reduce equation (2.2) to the form:

$$(Axdx + Bydy) + (Edx + Fdy) + C(xdy - ydx) = 0. \tag{2.4}$$

Unless $AB + C^2 = 0$, a suitable translation of the axes to a new origin will make E and F zero. When $AB + C^2 = 0$, while $C \neq 0$, we may make $E/F = \sqrt{-A/B}$ if we take the right sign for the radical. This gives as normal forms under *euclidean* transformations:

$$Axdx + Bydy + C(xdy - ydx) = 0, \tag{2.51}$$

$$(ax - by)(adx + bdy) + L(adx - bdy) = 0, \tag{2.52}$$

$$Bydy + Edx = 0. \tag{2.53}$$

By an affine transformation, we may further reduce equation

(2.51) to one or other of the forms:

$$(2.511) \qquad xdx + ydy + C(xdy - ydx) = 0,$$

$$(2.512) \qquad xdx - ydy + C(xdy - ydx) = 0,$$

$$(2.513) \qquad ydy + xdy - ydx = 0,$$

$$(2.514) \qquad xdy - ydx = 0,$$

$$(2.515) \qquad xdx + ydy = 0,$$

$$(2.516) \qquad xdx - ydy = 0.$$

Here C may be positive or negative, but is different from zero. To reduce equation (2.52) we introduce the lines $ax + by$ and $ax - by$ as new axes, and then change the scale to give:

$$(2.521) \qquad ydx - dy = 0,$$

or, when $L = 0$,

$$(2.522) \qquad ydx = 0.$$

Finally, from equation (2.53) we obtain:

$$(2.531) \qquad ydy - dx = 0,$$

$$(2.532) \qquad ydy = 0,$$

$$(2.533) \qquad dx = 0.$$

The integral curves of these eleven normal equations for *affine* transformations may be described geometrically. We have, in order, a family of equiangular spirals (2.511); imaginary transforms of these, i.e. curves cutting a family of hyperbolas under constant (hyperbolically measured) angle (2.512); a family of exponential curves $y = ke^{x/y}$ (2.513); the straight lines through the origin (2.514); concentric circles (2.515); equilateral hyperbolas (2.516); exponential curves $y = ke^x$, projectively equivalent to those of (2.513), (2.521); parabolas (2.531); parallel lines for (2.522), (2.532), (2.533).

For the forms (2.511) and (2.512), which may be considered the general case, as well as for forms (2.515) and (2.516), there are among the integral curves two straight lines, the common tangents (real or imaginary) for the family of conics associated with each type. Such straight line solutions for the more specialized cases will be treated later in the table preceding Theorem IV, where projective transformations are used, and the line at infinity is also shown to be a solution.

3. Straight line integral curves. We have just established

Theorem I. When the coefficients of the Pfaffian in two variables are each polynomials of the first degree, there are in general two finite straight line integral curves.

That this theorem is not true when the degree of the polynomial coefficients is higher than the first, may be seen by writing in place of equation (2.1):

$$A - Bp = 0, \tag{3.1}$$

and differentiating it to get:

$$A_x + A_y p - B_x p - B_y p^2 - Bp' = 0. \tag{3.2}$$

This shows that the integral curves will have a zero second derivative whenever the co-ordinates x and y satisfy

$$A^2 B_y + AB(B_x - A_y) - B^2 A_x = 0. \tag{3.3}$$

In the case of the theorem, the partial derivatives reduced to constants, and this relation was equivalent to two linear relations between A and B, each of which gave a straight line,

$$A - Bm = 0, \tag{3.4}$$

with

$$m^2 B_y + m(B_x - A_y) - A_x = 0. \tag{3.5}$$

But, for the straight line (3.4), the derivative p satisfies

$$pmB_y + mB_x - pA_y - A_x = 0, \tag{3.6}$$

which, compared with (3.5), necessitates $p = m$, unless we have

$A_y - B_y m = 0$, and this case is disposed of by interchanging the roles of x and y in the discussion.

For the higher degrees, the left member of (3.3) no longer factors in the general case, and even when it does give linear factors these need not have derivatives satisfying (3.1). For the nth degree case, (3.3) gives a locus of the $(3n-1)$st degree. Whenever an integral curve cuts this locus, it has a point of inflexion at the intersection.

4. Projective transformations and homogeneous co-ordinates. We note that a projective transformation of the plane will, in general, cause the degree of the coefficients of the Pfaffian (2.1) to be increased by unity. The inverse transformation illustrates that there are some special equations which may have their degree lowered by a suitably chosen projective transformation.

To further discuss this situation, we introduce homogeneous co-ordinates by putting:

$$x = x_1/x_3, \qquad y = x_2/x_3. \tag{4.1}$$

With this notation, a Pfaffian (2.1) whose coefficients are polynomials of the nth degree becomes, after multiplication by x_3^{n+2},

$$A_1(x_3dx_2 - x_2dx_3) + A_2(x_1dx_3 - x_3dx_1) = 0. \tag{4.2}$$

Here the coefficients A_1 and A_2 are homogeneous polynomials in the three variables, of the nth degree.

The equation just written suggests the more general form:

$$\begin{vmatrix} dx_1 & dx_2 & dx_3 \\ x_1 & x_2 & x_3 \\ A_1 & A_2 & A_3 \end{vmatrix} = 0, \tag{4.21}$$

of which it is a special case.

Another form is

$$P_1dx_1 + P_2dx_2 + P_3dx_3 = 0, \tag{4.22}$$

where the P_i are homogeneous functions such that

$$P_1x_1 + P_2x_2 + P_3x_3 = 0. \tag{4.23}$$

These forms are readily shown to be equivalent. For, when the determinant of (4.21) is multiplied out, it clearly gives a form like (4.22), with the condition (4.23) satisfied. Conversely, if any three homogeneous functions A_1, A_2, A_3 are taken such that

$$P_1A_1 + P_2A_2 + P_3A_3 = 0, \tag{4.24}$$

the equations (4.22), (4.23) and (4.24) will have as a consequence the equation (4.21). For example, we might use the identity

$$P_1(-P_3) + P_2(P_1) + P_3(0) = 0, \tag{4.25}$$

and so obtain the reduced form (4.2). This shows that all the three forms given are in one sense equivalent.

Consider next the behavior under projective transformations. The forms (2.1) and (4.2) are unchanged by affine transformations, but in general change their degree under projective transformations. The forms (4.22) and (4.23), on the other hand, are unaffected as to form and degree not only under projective transformations, but also under changes in the ratios of the co-ordinates of the type

$$x_1 = Kx_1', \qquad x_2 = Kx_2', \qquad x_3 = Kx_3', \tag{4.26}$$

the factor K being any function of the co-ordinates.

Since the A_i are any homogeneous functions satisfying (4.24), they are to some extent arbitrary, and even if we take them of minimum possible degree, we may replace them by $A_i - Lx_i$, where L is one degree lower than the A_i, as is seen from the form (4.21). If we begin with coefficients of the nth degree in the form (4.22), and reduce to the form (4.2) by means of (4.25) we shall have A_1 and A_2 of the nth degree. However, if we begin with coefficients A_1 and A_2 in (4.2) of the nth degree, and go over to form (4.22), the P_i will in general be of the $(n + 1)$st degree. Thus there are two cases of Pfaffians with

coefficients of the nth degree. Those reducible to the form (4.2) with A_i of the nth degree, and those reducible to the form (4.23) with P_i of the nth degree. Their typical equations in non-homogeneous co-ordinates are obtained by putting $x_3 = 1$, $dx_3 = 0$ in these forms, giving:

$$Adx + Bdy = 0, \tag{4.27}$$

$$Adx + Bdy + C(xdy - ydx) = 0. \tag{4.28}$$

For degrees higher than the first, these two cases are essentially different, but for constant or first degree coefficients, they are projectively equivalent. For, a projective transformation

$$x_i' = \Sigma a_{ij} x_j, \tag{4.3}$$

may be effected by noting that:

$$\begin{vmatrix} dx_1 & dx_2 & dx_3 \\ x_1 & x_2 & x_3 \\ A_1 & A_2 & A_3 \end{vmatrix} \cdot \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} \tag{4.31}$$

$$= \begin{vmatrix} \Sigma a_{1j} dx_j & \Sigma a_{2j} dx_j & \Sigma a_{3j} dx_j \\ \Sigma a_{1j}\, x_j & \Sigma a_{2j}\, x_j & \Sigma a_{3j}\, x_j \\ \Sigma a_{1j} A_j & \Sigma a_{2j} A_j & \Sigma a_{3j} A_j \end{vmatrix}.$$

In this all the sums are from 1 to 3. This equation shows that the equation (4.21) is equivalent to a similar equation in the primed letters, with

$$A_i' = \Sigma a_{ij} A_j, \tag{4.32}$$

at least when the determinant of the a_{ij} is not zero, i.e. when the transformation (4.3) is non-degenerate.

When the coefficients A_i are constant, we may find quantities a_{ij} such that we have simultaneously:

$$\begin{aligned} A_2' &= \Sigma a_{2j} A_j = 0, \\ A_3' &= \Sigma a_{3j} A_j = 0. \end{aligned} \tag{4.33}$$

On dividing out A_3', we reduce our form to:

$$x_3dx_2 - x_2dx_3 = 0, \tag{4.34}$$

which is the typical form for this case, and, in non-homogeneous co-ordinates, is of type (4.27).

When the coefficients are of the first degree, and not reducible to constants by division of a common linear factor, we write

$$A_i = \Sigma c_{ij}x_j. \tag{4.35}$$

We may now determine a_{ij} so that we have, simultaneously,

$$\begin{aligned} \Sigma a_{3j}c_{j1} &= ka_{31}, \\ \Sigma a_{3j}c_{j2} &= ka_{32}, \\ \Sigma a_{3j}c_{j3} &= ka_{33}, \end{aligned} \tag{4.36}$$

by taking k so that these equations are consistent. This makes

$$A_3' = kx_3', \tag{4.37}$$

so that by subtracting k times the second row from the third row, we may replace A_3' by 0, and reduce to a form which, in non-homogeneous co-ordinates, is of type (4.27).

As we have already classified the first degree case of the Pfaffian (4.27), we may at once write down projective normal forms for the form (4.28), which we have just shown is projectively equivalent to the earlier case. In this way we obtain as canonical values for A_1, A_2, A_3:

$$-x_2 + Cx_1, \qquad \pm x_1 - Cx_2, \qquad 0, \tag{4.41}$$

to which the forms (2.511) and (2.512) are reducible,

$$-x_2, \qquad \pm x_1, \qquad 0, \tag{4.42}$$

to which the forms (2.515) and (2.516) are reducible,

$$x_3, \qquad x_2, \qquad 0, \tag{4.43}$$

to which the types (2.513) and (2.521) are reducible,

$$x_2, \qquad x_3, \qquad 0, \tag{4.44}$$

for the type (2.531),

$$0, \quad x_2, \quad 0, \tag{4.45}$$

for the types (2.514) and (2.522), and

$$x_2, \quad 0, \quad 0 \tag{4.46}$$

for the types (2.532) and (2.533).

In this set, the plus or minus signs may be dispensed with if we use imaginary projective transformations.

These forms may also be obtained from a second point of view. For, since the first two rows of the determinant (4.21) transform in the same way, every reduction of the collineation

$$x_i' = A_i, \tag{4.5}$$

will give a reduction of our determinant. The canonical forms for such collineations not the identity, and non-singular are well known,[3] there being five types:

$$a_{11}x_1, \quad a_{22}x_2, \quad a_{33}x_3, \quad (a_{11}, a_{22}, a_{33} \text{ distinct}), \tag{4.51}$$

$$a_{11}x_1, \quad a_{22}x_2, \quad a_{32}x_2 + a_{22}x_3, \quad (a_{11} \neq a_{22}, a_{32} \neq 0), \tag{4.52}$$

$$x_1, \quad x_2 + a_{21}x_1, \quad x_3 + a_{31}x_1 + a_{32}x_2, \quad (a_{21} \neq 0, a_{32} \neq 0), \tag{4.53}$$

$$x_1, \quad x_2, \quad a_{33}x_3, \quad (a_{33} \neq 1), \tag{4.54}$$

$$x_1, \quad x_2, \quad x_3 + a_{32}x_2, \quad (a_{32} \neq 0), \tag{4.55}$$

where, in each case, we have written x_1', x_2', x_3' in order.

This shows that our A_i may be reduced to these forms. Then, a suitable multiple of the second row of the determinant (4.21) may be subtracted so as to lessen the number of terms. Finally, an appropriate combination of columns and change of notation will convert them to the standard forms (4.41) to (4.46). The additional form is due to the fact that when, in (4.51), $a_{11} = -a_{22}$, the C of (4.41) vanishes, and we have the form

[3] Cf., e.g., Woods, *Higher Geometry*, pp. 240–246, Veblen and Young, *Projective Geometry*, vol. I, pp. 274–276, and Bôcher, *Higher Algebra*, pp. 290–295.

(4.42). For this first case, the first two columns must be replaced by their sums and differences respectively, and

$$C = \frac{a_{22} + a_{11}}{a_{22} - a_{11}}. \tag{4.56}$$

The classification just made assumes that the collineation (4.5) is non-singular, i.e. has a non-vanishing determinant. When (4.5) as it stands has a zero determinant, we modify the last row of (4.21) by a multiple of the second, thus replacing (4.5) by

$$x_i' = A_i + kx_i. \tag{4.57}$$

This is singular for at most three values of k, we use any value distinct from these three. We have also omitted the identity, as in this case the A_i would be multiples of the x_i, and (4.21) would cease to be an equation.

Since all the collineations (4.57) have the same fixed points and fixed lines, it is natural to ask the significance of these for the Pfaffian. Recalling that a singular point of the Pfaffian is a point for which the ratios of the differentials are not restricted, so that all the differential elements at the point are solutions, we may state:

Theorem II. The singular points of a Pfaffian (4.21) *are the fixed points of the corresponding collineations* (4.57).

For, the singular points of the Pfaffian (4.21) are those for which the three minors of the first row vanish, while the fixed points of the collineation (4.57) are those for which the x_i and the A_i are proportional, and these are clearly equivalent conditions.

We also have:

Theorem III. The straight line integral curves of the Pfaffian (4.21) *are the fixed straight lines of the corresponding collineations* (4.57).

For, a fixed line of the collineation (4.57), not necessarily pointwise fixed, corresponds to a linear combination of the x_i equal to a constant times the corresponding linear combination

of the A_i. Consequently, if this linear combination is set equal to zero, the same will be true for this combination of the dx_i, and the A_i, and we shall have a solution of the Pfaffian (4.21). Conversely, if

$$\Sigma b_i x_i = 0, \tag{4.58}$$

is a straight line solution of the Pfaffian, for points and differentials on this line we have also:

$$\Sigma b_i dx_i = 0, \tag{4.59}$$

which show that the b_i are proportional to the coefficients of the A_i in (4.21), and since this vanishes,

$$\Sigma b_i A_i = 0. \tag{4.6}$$

Thus the line is a fixed line of the collineation.

These theorems, combined with the known fixed points and fixed lines of the collineations (4.51) to (4.55), and the methods of reduction to the forms (4.41) to (4.46) enable us to easily write down the singular points, and singular lines for these latter forms. We have, namely:

(4.41) fixed lines, $x_1 \pm ix_2 = 0$ (upper sign) or $x_1 \pm x_2$ (lower sign) together with $x_3 = 0$; the three intersections of these giving the fixed points,

(4.42) same as (4.41),

(4.43) fixed lines, $x_2 = 0$, $x_3 = 0$; the fixed points are (1, 0, 0) and (0, 1, 0),

(4.44) fixed line, $x_3 = 0$, fixed point (1, 0, 0),

(4.45) fixed lines, $x_2 = 0$ and all the lines through (0, 1, 0); fixed points (0, 1, 0), and all the points on $x_2 = 0$,

(4.46) fixed lines, all those through (1, 0, 0); fixed points, all the points on $x_2 = 0$.

The possibilities as to singular points and straight line solutions are given in detail above, but we note one evident property in

Theorem IV. Every straight line integral curve of a Pfaffian (4.21) *goes through at least one singular point, and each singular point is on at least one straight line solution.*

5. The three-dimensional affine case. While the Pfaffian

$$A(x, y, z)dx + B(x, y, z)dy + C(x, y, z)dz = 0, \tag{5.1}$$

does not in general possess an integrating factor, some of the methods of section 2 may be used in case the A, B and C are linear functions. We will omit from our discussion all the integrable cases; in particular this excludes the case where the coefficients A, B, C are all constants. We note that terms like

$$axdx, \qquad bdx, \tag{5.11}$$

are integrable as they stand, while the remaining six terms may be paired up, and reduced by the relation

$$cydx+exdy=(c+e)(ydx+xdy)/2+(c-e)(ydx-xdy)/2. \tag{5.12}$$

Thus the Pfaffian may be written in the form:

$$dQ + a(ydz - zdy) + b(zdx - xdz) + c(xdy - ydx) = 0, \tag{5.13}$$

where dQ is the exact differential of a quadratic expression in x, y, z.

Let us determine those planes

$$z = px + qy + r, \tag{5.14}$$

in which the integral curves are a family of conics which form the solution of an equation:

$$dR(x, y) = 0. \tag{5.15}$$

Since the form of this equation is unchanged by affine transformations, e.g. that used to go from co-ordinates in the plane (5.14) to the x, y of space; elimination of z from (5.14) and (5.13) will have to give such an equation. That is, the inexact part of the resulting equation:

$$(ap + bq - c)(ydx - xdy) \tag{5.16}$$

must vanish identically, so that

$$c = pa + qb, \tag{5.17}$$

and the planes (5.14) sought are those parallel to the direction a, b, c. As this direction has a geometric interpretation, if we take a new z axis in this direction, the transformed Pfaffian (5.13) will have $a = b = 0$, and reduce to

$$dQ + C(xdy - ydx) = dQ + Cx^2 d(y/x) = 0. \tag{5.18}$$

The second form provides a simple illustration of a general theorem on Pfaffians in three variables, which states that they may be reduced to one of the three forms $dz + udv$, udv, or dz. The reducing transformation, usually a complicated functional transformation, in our case may be taken to be an orthogonal one. The last two types occur here when $dQ = 0$, or when in (5.13) $a = b = c = 0$. These are integrable cases, the integral surfaces being the planes through a fixed line in the first case and a family of quadrics or planes, in the second, with equation $Q = k$.

As the interpretation of the direction with x, y, z components a, b, c applies to oblique co-ordinates, or affine transformations of our space, the form (5.18) will be preserved as long as the direction of the z axis is unchanged. If this direction does not give an infinite intersection with the quadrics $Q = k$, we may take the corresponding diametral plane for these quadrics as the x, y plane, and reduce the Pfaffian to the form:

$$zdz + dR(x, y) + C(xdy - ydx) = 0. \tag{5.2}$$

When there is one infinite intersection, the Pfaffian (5.18) will be in the form:

$$dz + dR(x, y) + C(xdy - ydx), \tag{5.3}$$

or else will not contain z, and we shall have an integrable case.

We may now simplify the terms independent of z by affine transformations in the x, y plane previously used to reduce equation (2.2). By referring to the forms (2.511) to (2.533),

and taking account of the restrictions as to sign caused by the z term, after omitting the integrable cases, we find as canonical forms for the Pfaffian (5.13) under affine transformations:

$$(5.21)\qquad xdx + ydy + zdz + C(xdy - ydx) = 0,$$

$$(5.22)\qquad xdx + ydy - zdz + C(xdy - ydx) = 0,$$

$$(5.23)\qquad xdx - ydy + zdz + C(xdy - ydx) = 0,$$

$$(5.24)\qquad ydy + zdz + xdy - ydx = 0,$$

$$(5.25)\qquad ydy - zdz + xdy - ydx = 0,$$

$$(5.26)\qquad zdz + xdy - ydx = 0,$$

$$(5.27)\qquad zdz + ydx - dy = 0,$$

$$(5.28)\qquad zdz + ydx = 0,$$

$$(5.31)\qquad xdx + ydy + dz + C(xdy - ydx) = 0,$$

$$(5.32)\qquad xdx - ydy + dz + C(xdy - ydx) = 0,$$

$$(5.33)\qquad ydy + dz + xdy - ydx = 0,$$

$$(5.34)\qquad dz + xdy - ydx = 0,$$

$$(5.35)\qquad dz + ydx - dy = 0,$$

$$(5.36)\qquad dz + ydx = 0.$$

A picture of the configuration of differential elements satisfying any one of these may be visualized by noting that in planes through the z axis the integral curves are conics (or parallel lines for (5.26) and (5.34)), while in planes parallel to the x, y plane, the integral curves are the solutions of the forms (2.511) to (2.533) which have been described earlier.

6. Straight line integral curves. In view of the classification just made, we may regard

$$(6.1)\qquad axdx + bydy + czdz + C(xdy - ydx) = 0,$$

as the general type of Pfaffian, and if we add the Pfaffian

$$(6.2)\qquad axdx + bydy + edz + C(xdy - ydx) = 0,$$

we will have included all but the special cases (5.27), (5.28), (5.35) and (5.36). The straight line

$$(6) \qquad x = X + p_1 t, \qquad y = Y + p_2 t, \qquad z = Z + p_3 t,$$

will be an integral curve provided:

$$(6.11) \qquad ap_1^2 + bp_2^2 + cp_3^2 = 0,$$
$$ap_1 X + bp_2 Y + cp_3 Z + C(p_2 X - p_1 Y) = 0,$$

or,

$$(6.21) \qquad ap_1^2 + bp_2^2 = 0,$$
$$ap_1 X + bp_2 Y + ep_3 + C(p_2 X - p_1 Y) = 0.$$

For each value of X, Y, Z, except the singular point 0, 0, 0 for (6.11), these conditions give a linear and a quadratic relation on the p_i, so that they determine two values of their ratios. This establishes:

Theorem V. When the coefficients of the Pfaffian in three variables are each polynomials of the first degree, there are in general just two straight lines through each finite point with the exception of the singular point.

The integral curves of a Pfaffian of the present type, in any plane not parallel to the z-axis, will be those of a Pfaffian in two variables, and hence, by Theorem I, will include two straight lines. This establishes:

Theorem VI. For the Pfaffian (5.13), *there are in general, in each finite plane, not parallel to the axis, just two straight lines.*

This result is also evident from the fact that if

$$(6.3) \qquad Lx + My + Nz + K = 0,$$

is a given plane, and we put

$$(6.31) \qquad \frac{L}{aX - CY} = \frac{M}{bY + CX} = \frac{N}{cZ} \text{ or } \frac{N}{e},$$

these equations determine either the ratios of X, Y, Z, or the values of X and Y, and the factor of proportionality, or the value

of Z, may then be determined from (6.3). The two lines found from (6.11) or (6.21) for this point X, Y, Z will then be parallel to the plane (6.3) and, since they contain a point in it, will lie in the plane.

Incidentally, we note that the Pfaffian gives a plane through each point, and also a point in each plane.

It is interesting to note the way in which the straight lines just found may be grouped into the rulings of quadrics. Consider the quadrics:

$$ax^2 + by^2 + c(z - sC)^2 + abcs^2 = 0. \tag{6.4}$$

This will contain the straight line (6) as a ruling if:

$$\begin{aligned} ap_1^2 + bp_2^2 + cp_3^2 &= 0, \\ ap_1X + bp_2Y + cp_3(Z - sC) &= 0, \\ aX^2 + bY^2 + c(Z - sC)^2 + abcs^2 &= 0. \end{aligned} \tag{6.41}$$

Elimination of Z from the last two equations gives:

$$acp_3^2X^2 + bcp_3^2Y^2 + (ap_1X + bp_2Y)^2 + abc^2p_3^2s^2 = 0, \tag{6.42}$$

which, by means of the first equation, reduces to:

$$-ab(p_2X - p_1Y)^2 + abc^2p_3^2s^2, \tag{6.43}$$

or,

$$cp_3s = \pm (p_2X - p_1Y). \tag{6.44}$$

If the lower sign be taken, in view of (6.44) the second of the equations (6.41) becomes identical with the second of (6.11), and as the first of each set are identical, we see that:

One set of rulings of the family of quadrics (6.4) *are integral curves of the Pfaffian* (6.1); *the other set correspond to a similar Pfaffian with the sign of C reversed.*

For the lines on the Pfaffian (6.2), we start with the quadrics:

$$ax^2 + by^2 + 2ez/(1 + C/\sqrt{-ab}) + u = 0. \tag{6.5}$$

This will contain the straight line (6) as a ruling if:

$$(6.51)\qquad \begin{aligned} ap_1^2 + bp_2^2 &= 0,\\ ap_1X + bp_2Y + ep_3/(1 + C/\sqrt{-ab}) &= 0,\\ aX^2 + bY^2 + 2eZ/(1 + C/\sqrt{-ab}) + u &= 0. \end{aligned}$$

The first two of these equations may be combined to give:

$$(6.52)\qquad -ab(p_2X - p_1Y)^2 = p_3^2e^2/(1 + C/\sqrt{-ab})^2,$$

so that:

$$(6.53)\qquad (p_2X - p_1Y) = \pm p_3e/(\sqrt{-ab} + C),$$

and, if the lower sign be taken,

$$(6.54)\qquad ep_3 + C(p_2X - p_1Y) = p_3e/(1 + C/\sqrt{-ab}).$$

In consequence of this relation, the second of equations (6.51) becomes identical with the second of (6.21), and as the first of each set are the same, we conclude that:

One set of rulings of the family of quadrics (6.5) *are integral curves of the Pfaffian* (6.2), *the other set correspond to a similar Pfaffian with the sign of C reversed.*

Since changing s in the equation (6.4) converts one quadric into a second similar to it, with the origin as the center of similitude, this family of quadrics all have the same tangent cone with vertex at the origin. They also have the same intersection with the plane at infinity, and thus may be characterized as:

A family of quadrics with a common plane section (at infinity) and a common tangent cone.

We note that the quadrics are determined as soon as the vertex of the cone, the plane of section, and one of them is given.

To interpret the family of quadrics, (6.5), we observe that it is a limiting form of the family just discussed. For, if we set

$$(6.55)\qquad z = z' + e/c$$

in the Pfaffian (6.1), and let c approach zero, it takes the form

(6.2). If we make this change in equation (6.4) and at the same time put

$$s = \frac{e}{c(C + \sqrt{-ab})} - \frac{u}{2e\sqrt{-ab}}, \tag{6.56}$$

when c approaches zero, it assumes the form (6.5). This family of quadrics is such that each member is tangent to the plane at infinity at the same point, at which it osculates any other member of the family to the second order. That is, the common cone and plane section coincide, to form a tangent plane.

The discussion just made enables us to describe the rulings of the Pfaffians (5.21), (5.22), (5.23), as well as those of (5.31) and (5.32), as geometric configurations. For the remaining cases, the equations (6.4) and (6.5) no longer apply, owing to the vanishing of some of the constants, and the rulings do not group together to form quadrics in this way. The geometric interpretation of these cases may be found from consideration of the equations (6.11), (6.21) or those formed similarly to them. For (5.24) and (5.25) we have in each case two sets of lines, the lines of each set being parallel to a given plane ($y \pm iz = 0$, $y \pm z = 0$), and intersecting a given line ($y = 0$, and $x \pm iz = 0$, $x \pm z = 0$). Thus we have a second order congruence, consisting of two first order congruences. Each of these, considered projectively, is the set of all straight lines cutting two given straight lines.[4] For (5.26) we have just one such congruence, the rulings being all parallel to the plane $z = 0$, and intersecting the z-axis. For (5.27) the rulings are the lines tangent at finite points to the parabolic cylinder $z^2 + 4y = 0$, and parallel to the elements of the cone $z^2 + xy = 0$, together with lines in the xz plane parallel to the x axis. We notice that these quadrics are tangent at infinity, so that there are only two finite straight lines in each plane, and through each finite point, in general. For

[4] For the relation of the congruences here described to the general classification of congruences, compare Pascal's *Repertorium*, vol. II, part 2, edited by Timerding, Berlin, 1922, p. 1032, p. 1036 f.; R. Sturm, *Liniengeometrie*, vol. II, Leipzig, 1893, p. 23 f., p. 316 f.

(5.28) the rulings consist of all the lines parallel to the x axis, together with all the rulings of the family of cones $(z - z_0)^2 + xy = 0$. For (5.33) we have the set of rulings $y = \text{const.}$ on the hyperbolic paraboloids $z - z_0 = xy$. For (5.34) we have the general linear complex, or null system. For (5.35) and (5.36) one set of rulings consists of all the lines parallel to the y axis, while the other set are the rulings $y = \text{const.}$ on a family of hyperbolic paraboloids ($z - z_0 + xy - x = 0$, $z - z_0 + xy = 0$ respectively).

We may obtain a further insight into the structure of the quadratic congruences (6.41) and (6.51) by considering them as the intersection of two quadratic complexes. For this discussion it is desirable to introduce the Plücker line coördinates.[5] The coördinates of a line joining the point x_1, y_1, z_1 with x_2, y_2, z_2 may be taken as proportional to

$$
\begin{aligned}
(6.6) \qquad & rq_1 = x_2 - x_1, \qquad rq_4 = y_1z_2 - y_2z_1, \\
& rq_2 = y_2 - y_1, \qquad rq_5 = z_1x_2 - z_2x_1, \\
& rq_3 = z_2 - z_1, \qquad rq_6 = x_1y_2 - x_2y_1,
\end{aligned}
$$

where r is an arbitrary factor of proportionality.

For the line defined by (6), on taking $r = t_2 - t_1$, the difference of the parameters for the two points, we find:

$$
\begin{aligned}
(6.61) \qquad & q_1 = p_1, \qquad q_4 = Yp_3 - Zp_2, \\
& q_2 = p_2, \qquad q_5 = Zp_1 - Xp_3, \\
& q_3 = p_3, \qquad q_6 = Xp_2 - Yp_1.
\end{aligned}
$$

We recall the identical relation:

$$(6.62) \qquad q_1q_4 + q_2q_5 + q_3q_6 = 0.$$

In these coördinates, the first of equations (6.11) becomes:

$$(6.63) \qquad aq_1^2 + bq_2^2 + cq_3^2 = 0.$$

To transform the second of equations (6.11), we note that by (6.61)

[5] Pascal's *Repertorium*, l.c., vol. II, part 2, p. 990, p. 1009.

$$(6.64) \qquad Z = (q_5 + Xq_3)/q_1, \qquad Y = (-q_4 + Xq_3)/q_1,$$

and when these are inserted in the second of (6.11), the coefficient of X is zero by (6.63) and the first three relations (6.61), leading to the transformed equation:

$$(6.65) \qquad cq_3q_5 - bq_2q_6 + Cq_6q_1 = 0.$$

Had we used Y or Z in place of X, the transformed equation would have been:

$$(6.66) \qquad aq_1q_6 - cq_3q_4 + Cq_6q_2 = 0,$$

or

$$(6.67) \qquad bq_2q_4 - aq_1q_5 + Cq_6q_3 = 0.$$

These last three equations are readily seen to be equivalent in view of (6.62) and (6.63), since on multiplying them in turn by q_4, q_5, q_6 respectively, and summing, we obtain, essentially, (6.62); while on multiplying by aq_1, bq_2, cq_3 respectively and summing, we obtain, essentially, (6.63). Thus the congruence we started with is given as the intersection of the complex (6.63), and any one of the other three, or as the intersection of any two of the latter. A more symmetric complex, which, adjoined with (6.63) gives the congruence, is found by eliminating q_1, q_2 and q_3 from the latter set of four equations. The result is:

$$(6.68) \qquad bcq_4^2 + acq_5^2 + (ab + C^2)q_6^2 = 0.$$

The two complexes (6.63) and (6.68) are each singular, since each involves only one triplet of coördinates. The condition for a line of the complex $F(q_1, q_2, q_3, q_4, q_5, q_6) = 0$ to be singular is that

$$(6.69) \qquad \partial F/\partial q_1 \cdot \partial F/\partial q_4 + \partial F/\partial q_2 \cdot \partial F/\partial q_5 + \partial F/\partial q_3 \cdot \partial F/\partial q_6 = 0$$

for the coördinates of this line, and the complex is singular when every line belonging to it is singular.[5] In this case the complex is known to be formed either of the lines cutting a curve, or tangent to a surface. The complex (6.63) is of the first type,

the curve being the conic at infinity, while the complex (6.68) is of the second type, the lines being all tangent to a certain quadric cone with vertex at the origin. The equation of this cone is

$$aq_1^2 + bq_2^2 + abcq_3^2/(C^2 + ab) = 0. \tag{6.7}$$

This is readily seen to be the common tangent cone for the family of quadrics (6.4).

We may by similar methods transform the congruence (6.21) to Plücker coördinates. We thus obtain, as the equations of complexes containing the congruence:

$$aq_1^2 + bq_2^2 = 0, \tag{6.71}$$

$$eq_3q_1 - bq_2q_6 + Cq_1q_6 = 0, \tag{6.72}$$

$$eq_3q_2 + aq_1q_6 + Cq_2q_6 = 0, \tag{6.73}$$

$$eq_3^2 + bq_2q_4 - aq_1q_5 + Cq_3q_6 = 0. \tag{6.74}$$

On eliminating q_1, q_2 from these, we find:

$$eq_3 + (C + \sqrt{-ab})q_6 = 0. \tag{6.75}$$

This is in agreement with equation (6.53).

We have seen that the straight line curves may be determined from the Pfaffian. It is perhaps worth pointing out that, except in one or two very specialized cases, the Pfaffian equation is determined by the straight lines. For, there are in general two such straight lines through each point, and these determine the tangent plane to the curves in the Pfaffian, i.e. the Pfaffian itself, to within a factor.

7. Straight lines of the general Pfaffian. Let us consider a Pfaffian (5.1) in which the coefficients are *not* linear functions, and investigate the situation as to straight line integral curves. The condition that (6) satisfy (5.1) is:

$$Ap_1dt + Bp_2dt + Cp_3dt = 0, \tag{7.1}$$

where the variables in the coefficients are $X + p_1t$, etc. On

expanding this relation in powers of t, we find the conditions:

$$(7.2) \qquad \begin{gathered} Ap_1 + Bp_2 + Cp_3 = 0, \\ \sum_{i,j=1}^{3} A_j^i p_i p_j = 0, \\ \sum_{i,j,k=1}^{3} A_{jk}^i p_i p_j p_k = 0, \text{ etc.}, \end{gathered}$$

where the superscripts running from 1 to 3 give A, B, C in succession, and the subscripts denote partial differentiation with respect to x, y, z in succession. The variables in the coefficients are here X, Y, Z. This series gives an infinite number of necessary conditions for the functions at each point, and no finite number of these conditions are sufficient, i.e. these conditions are, in general independent. This is proved by a special example:

$$(7.21) \qquad x^n dx + y dy + dz = 0.$$

This expression is integrable, and the integral surfaces are

$$(7.22) \qquad x^{n+1}/(n+1) + y^2/2 + z = \text{const.}$$

As this family of surfaces contains all the integral curves of the Pfaffian, but no straight lines through a general point, for $n > 1$, there are no straight line integral curves for the Pfaffian (7.21). However, the straight line:

$$(7.23) \qquad x = t, \quad y = Y, \quad z = Z,$$

for which

$$(7.24) \qquad X_t = 0, \quad p_1 = 1, \quad p_2 = 0, \quad p_3 = 0,$$

satisfies the first $(n-1)$ conditions (7.2) and only fails to satisfy the nth condition.

Evidently, unless very special circumstances are at hand, like linearity of the coefficients, there will be no straight lines through a general point satisfying a given Pfaffian. We may, however, have any number of straight lines for a particular point. An example of this is the Pfaffian:

$$(7.25) \qquad e^z dz + x^n dx + x^n f(y/x) dy = 0,$$

where $f(u)$ is an nth degree polynomial in u, such that

$$1 + uf(u) = 0 \tag{7.26}$$

has $(n + 1)$ distinct roots, $u_1, u_2, \ldots, u_{n+1}$. At a general point, the equations (7.2) will be inconsistent. However, there are through the origin exactly $(n + 1)$ straight lines, namely:

$$z = 0, \qquad y = u_i x, \qquad (i = 1, 2, \ldots, n + 1). \tag{7.27}$$

8. The three-dimensional projective case. As in section 4, we find that the set of Pfaffians (5.1) with coefficients of a given degree are not transformed into the same set by projective transformations. By analogy with equations (4.21) and (4.28) we are led to construct the more generalized form:

$$A(ydz - zdy) + B(zdx - xdz) + C(xdy - ydx) + Fdx + Gdy + Hdz = 0. \tag{8.1}$$

The Pfaffians of this type with coefficients of a given degree are transformed into Pfaffians of the same type by projective transformations. This is seen by introducing the homogeneous coördinates:

$$x = x_1/x_4, \qquad y = x_2/x_4, \qquad z = x_3/x_4. \tag{8.2}$$

For these coördinates, the form (8.1) is

$$A(x_2dx_3 - x_3dx_2) + B(x_3dx_1 - x_1dx_3) + C(x_1dx_2 - x_2dx_1) + F(x_4dx_1 - x_1dx_4) + G(x_4dx_2 - x_2dx_4) + H(x_4dx_3 - x_3dx_4) = 0. \tag{8.11}$$

Since a projective transformation is a linear transformation of these coördinates, it clearly leaves the type of form unchanged. The coefficients may here be taken as homogeneous polynomials of the same degree (n).

When we take them to be of degree unity, or linear expressions, each coefficient in the forms (8.1) and (8.11) has four constants, so that there are apparently 24 constants involved, less one which can be divided out. However, the relations between the expressions in parentheses reduce the number of

significant constants to 19. Thus, for example, in (8.1) we may assume the constant terms of A, B and C removed to F, G, H, and place one restriction on the distribution of the constants in these first three coefficients, since increasing A, B, C by mx, my, mz respectively has no effect on the form.

The discussion of the two variable case, and the forms (4.21) and (4.22) suggest equivalent forms for (8.11), which we rewrite as:

$$(8.12)\quad dx_1(Bx_3 - Cx_2 + Fx_4) + dx_2(-Ax_3 + Cx_1 + Gx_4) + dx_3(Ax_2 - Bx_1 + Hx_4) + dx_4(-Fx_1 - Gx_2 - Hx_3) = 0.$$

This is a form

$$(8.13)\qquad P_1dx_1 + P_2dx_2 + P_3dx_3 + P_4dx_4 = 0,$$

in which the P_i are homogeneous expressions of the $(n+1)$st degree, satisfying the condition:

$$(8.14)\qquad P_1x_1 + P_2x_2 + P_3x_3 + P_4x_4 = 0.$$

If, now, the P_i in (8.13) are arbitrarily chosen quadratic expressions, satisfying (8.14), the number of independent constants is 19, i.e. 40 in (8.13), less 20 for the relations of (8.14) less one to be divided out. Thus in the linear case these forms (8.11) and (8.13) appear to be equivalent. We readily check this for the nth degree case by noting that if we use the abbreviations:

$$(8.15)\qquad P_1(0, x_2, x_3, x_4) = P_1',\quad P_1(0, 0, x_3, x_4) = P_1'', \text{ etc.}$$

and put first x_1, and then x_1 and x_2 zero in the identical relation (8.14), we have:

$$(8.16)\qquad P_2'x_2 + P_3'x_3 + P_4'x_4 = 0,\quad P_3''x_3 + P_4''x_4 = 0.$$

This last equation shows that

$$(8.17)\qquad H = P_3''/x_4 = -P_4''/x_3$$

is an nth degree expression, since the $(n+1)$st degree numera-

tors vanish when x_4 and x_3 vanish respectively, and hence contain these variables as factors. We next define

$$(8.18) \qquad A = \frac{P_3' - P_3''}{x_2}, \qquad G = \frac{-P_4' + P_4''}{x_2},$$

which are again nth degree expressions, since the numerators vanish with x_2. Finally, we put

$$(8.19) \quad B = \frac{-P_3 + P_3'}{x_1}, \quad C = \frac{P_2 - P_2'}{x_1}, \quad F = \frac{-P_4 + P_4'}{x_1},$$

which are seen to be of the nth degree by the same argument. If we now insert the values given for A, B, C, F, G, H in equation (8.12), the coefficients of dx_3 and dx_4 are seen to be P_3 and P_4, directly, while the coefficients of dx_1 and dx_2 are equal to P_1 and P_2 because of the relations (8.14) and (8.16) connecting the P_i, P_i' and P_i''.

We may also write a determinant form:

$$(8.21) \qquad \begin{vmatrix} dx_1 & dx_2 & dx_3 & dx_4 \\ x_1 & x_2 & x_3 & x_4 \\ A_1 & A_2 & A_3 & A_4 \\ B_1 & B_2 & B_3 & B_4 \end{vmatrix} = 0.$$

If the A_i are homogeneous polynomials of the mth degree, and the B_i are homogeneous polynomials of the m'th degree, when this determinant is expanded it will give an expression of the form (8.13) with P_i homogeneous polynomials of the $(m + m')$th degree, satisfying the relation (8.14). Conversely, when the form (8.13) is given, we have merely to select two sets of functions A_i and B_i such that:

$$(8.22) \qquad \begin{aligned} P_1A_1 + P_2A_2 + P_3A_3 + P_4A_4 &= 0, \\ P_1B_1 + P_2B_2 + P_3B_3 + P_4B_4 &= 0, \end{aligned}$$

in order to obtain a form (8.21) whose vanishing is a consequence of equation (8.13). However, the degree $(m + m')$ will in general be greater than n, the determinant containing a factor of

the $(m + m' - n)$th degree. Moreover, it is frequently impossible to select A_i and B_i for which no such factor cancels out. For example, the Pfaffian

$$x_1dx_2 - x_2dx_1 + x_3dx_4 - x_4dx_3 = 0, \tag{8.23}$$

implies

$$\begin{vmatrix} dx_1 & dx_2 & dx_3 & dx_4 \\ x_1 & x_2 & x_3 & x_4 \\ x_1 & x_2 & 0 & 0 \\ x_4 & 0 & x_3 & 0 \end{vmatrix} = 0. \tag{8.24}$$

The additional factor is here x_3x_4, of the second degree. But, every determinant which gives (8.23) must contain an additional factor of the second degree, at least. For, if not, one of the last two rows would have to be of less than the first degree, and hence consist of constants. In this case, a linear combination of columns, with a corresponding linear (projective) transformation of the variables, would yield a new form with three zeros in this row, and hence give a Pfaffian with only three of the differentials present. This is impossible, as (8.23), which represents the general linear complex, or null system, involves all four differentials no matter what coördinates are used. The appearance of this extraneous factor makes the determinant form difficult to work with when classifying forms for which the degree of the P_i are fixed, and we shall make no further use of (8.21) for this reason.

Let us now return to the form (8.13). When the P_i are constants, we have an integrable case. When they are of the first degree, the coefficients of the corresponding form (8.11) are constants, and since the equation then becomes a linear combination of the Plücker coördinates, (6.6), for the straight line through x_i with direction dx_i, we have a linear complex. By a suitable linear transformation of coördinates, it may be reduced to the canonical form:

$$x_1dx_2 - x_2dx_1 = 0, \tag{8.25}$$

if a special type of linear complex, and to the form (8.23) if a general type of linear complex.

When the P_i are quadratic, the coefficients of (8.11) are of the first degree. If it is possible to make A, B and C zero by a proper change of coördinates, we may by a corresponding choice of the plane at infinity for the non-homogeneous coördinates reduce the Pfaffian to the form (5.1) which has been classified in the table (5.21) through (5.36). If no such choice of the plane at infinity is possible, we take it as fixed, for the present, at random and change to the non-homogeneous Cartesian coordinates. We may still apply affine transformations, which will effect projective transformations in the plane at infinity. If we remove the constant terms from A, B, C, to F, G, H, the form

$$(8.27)\quad A(ydz - zdy) + B(zdx - xdz) + C(xdy - ydx)$$

will be homogeneous, and is analogous to (4.21). This latter form has been classified in equations (4.41) through (4.46). By finding the two-dimensional projective transformation which effects this reduction, and now interpreting it as a three-dimensional affine transformation, we may leave the general form of the Pfaffian (8.11) unchanged, but reduce the highest terms to canonical form. We thus obtain as a first stage of classification, the forms:

$$(8.31)\quad (-y + Cx)(ydz - zdy) + (\pm x - Cy)(zdx - xdz) + Fdx + Gdy + Hdz = 0,$$

$$(8.32)\quad -y(ydz - zdy) \pm x(zdx - xdz) + Fdx + Gdy + Hdz = 0,$$

$$(8.33)\quad z(ydz - zdy) + y(zdx - xdz) + Fdx + Gdy + Hdz = 0,$$

$$(8.34)\quad y(ydz - zdy) + z(zdx - xdz) + Fdx + Gdy + Hdz = 0,$$

$$(8.35)\quad y(zdx - xdz) + Fdx + Gdy + Hdz = 0,$$

$$(8.36)\quad y(ydz - zdy) + Fdx + Gdy + Hdz = 0.$$

In these forms F, G, H are linear expressions in x, y, z. In general, a suitable translation will free them of the constant term, though this may be impossible for all three in certain specialized cases. From the discussion of (4.51) and (4.56) we note that

$$(8.315)\quad x(ydz - zdy) + ky(xdz - zdx) + Fdx + Gdy + Hdz = 0$$

is an equivalent form for (8.31) when $k \neq -1$, and (8.32) when $k = -1$, with both signs if imaginary transformations are employed.

The six forms first given include all possible types. Moreover, there are forms which can only be classified as one of these. However, without restrictions on F, G, H, the types are not mutually exclusive since a projective transformation, leading to a new choice of the plane at infinity, will sometimes take a form from one class to another. This ambiguity may be avoided by regarding the types in order of simplicity as written, and in each case so take the plane at infinity as to give the simplest type. We shall not enter into the conditions on F, G, H, imposed by this convention.

As the sub-division of the six cases analogous to our earlier classifications would lead to an enormous number of special types, we shall not attempt it, but instead will indicate how any one type can be treated by studying a particular example of (8.31), the " general " case, namely,

$$(8.41)\quad x(ydz - zdy) + y(xdz - zdx) + xdx + ydy - 2zdz = 0.$$

This is written in the alternative form (8.315). The straight line (6) will be an integral curve of this Pfaffian, provided:

$$(8.42)\quad p_1p_3Y + p_2p_3X - 2p_1p_2Z + p_1^2 + p_2^2 - 2p_3^2 = 0,$$
$$(8.43)\quad p_1(X - YZ) + p_2(Y - XZ) + 2p_3(XY - Z) = 0.$$

When the point X, Y, Z is not on all three of the quadrics:

$$(8.44)\qquad xy = z, \qquad xz = y, \qquad yz = x,$$

the equations determine in general two values of the ratios

$p_1 : p_2 : p_3$, so that there are just two straight lines through the point.

The singular points on all three of the quadrics (8.44) are found to be:

(8.45) $(0,0,0), (1,1,1), (1,-1,-1), (-1,1,-1), (-1,-1,1).$

Through each of these points there is a quadric cone of straight lines. Thus these are second degree singular points. Besides these, there are certain other points which, while not singular for the Pfaffian (8.41) since the tangent plane is definitely determined, are singular for the congruence of straight line integral curves, since every straight line in the tangent plane is an integral curve. These first degree points are:

(8.46) $(-1,-1,-1), (-1,1,1), (1,-1,1), (1,1,-1).$

Incidentally, we notice that the congruence of straight lines has just three in a general plane, since the equation in p_1/p_2 obtained by putting $Z = AX + BY + C$, $p_3 = Ap_1 + Bp_2$, $Y = 0$ is of the third degree. Placing $Y = 0$ involves no loss of generality, since from the form of equations (8.42) and (8.43) as well as their geometric interpretation, if X, Y, Z, p_i satisfy them, the same will be true of $X + p_1 t$, $Y + p_2 t$, $Z + p_3 t$, p_i.

We have only given five solutions of the three quadrics (8.44), while eight were to be expected. The other three are at infinity. They have *no* geometric significance, since when we find their homogeneous coördinates, and substitute them in the Pfaffian for these coördinates, the coefficient of dx_4 which must then be introduced does not vanish.

The situation we have just explored for the Pfaffian (8.41) is typical of that for the general forms (8.31) through (8.36), when the F, G, H do not satisfy special conditions. That is, there are two straight lines through the general point which are integral curves, a certain number of points, singular for the congruence, for which a plane pencil of lines exists, and some second degree points, singular for the Pfaffian as well as for the congruence.

These last lie on the three quadrics:

$$(8.5) \qquad Bz - Cy + F = 0, \qquad -Az + Cx + G = 0, \qquad Ay - Bx + H = 0.$$

The notation goes back to (8.11), (8.12) and the equations may be written explicitly as to the second degree terms by comparing (8.27) with the set (8.31) through (8.36). For the first four cases we have three quadrics, with five finite intersections, giving singular points of the Pfaffian, and three intersections at infinity. For the last two cases we have two quadrics and a plane, with three finite intersections, giving singular points of the Pfaffian, and one intersection at infinity. The intersections at infinity only have a geometric interpretation when the further condition:

$$(8.51) \qquad Fx + Gy + Hz = 0,$$

is satisfied. The form of (8.12) suggests that this be adjoined to the set (8.5). This is satisfied in consequence of (8.5) when all the coördinates are finite, but not otherwise. When (8.51) holds, the quadrics (8.5) have a curve of intersection, and there is a singular line for the Pfaffian. The cases which specialized into (5.1) and the forms (5.21) through (5.36) are for the most part an example of this, the quadrics (8.5) here intersecting in a conic and a point. The plane of the conic was already at infinity, and the reduction we used brought the point to the origin. There are other cases, however, where the singular line is a third degree curve which degenerates into a conic and a straight line.

9. The n-dimensional case. The methods we have used may be applied to the Pfaffian in n-variables,

$$(9.1) \qquad \Sigma A_i dx_i = 0. \qquad (i = 1, 2, \ldots, n)$$

though the details become in some cases rather complicated. In particular, we note that when the A_i are expressions of the first degree, the form

$$(9.11) \qquad \Sigma a_{ij} x_j dx_j + \Sigma b_i dx_i = 0, \qquad (i, j, = 1, 2, \ldots, n)$$

contains a quadratic cone (in $n-1$ dimensions) of straight line integral curves, through each point not in certain special classes.

This is also true for the form:

$$(9.2)\quad \Sigma A_{ij}(x_i dx_j - x_j dx_i) = 0, \qquad (i, j = 1, 2, \ldots, n+1),$$

when the A_{ij} are homogeneous of the first degree. In this case, the form is explicitly:

$$(9.21)\qquad \Sigma a_{ijk} x_i x_j dx_k = 0,$$
$$a_{ijk} + a_{jki} + a_{kij} = 0, \qquad a_{ijk} = a_{jik},$$

all three subscripts running from 1 to $n+1$. The form is suggested as the projective generalization of (9.11).

THE CALCULATION OF RESISTANCES TO GROUND AND OF CAPACITANCE

By Herbert Bristol Dwight*

Synopsis.—The calculation of the resistance to ground of different shapes of conductors or groups of conductors is fundamentally the same as the calculation of capacitance. In this paper are presented (1) a series calculation for a cylinder of finite length; (2) curves for large groups of vertical ground rods; (3) a formula (No. 27) for buried metallic strip; (4) a formula (No. 31) for the capacitance of a round plate by the average potential method, showing that that method is in error by 8 per cent in that case; (5) a curve (Fig. 8) for the capacitance of a rectangular plate, with a correction for the error due to the average potential method. It is shown that the error due to the average potential method is very small in the case of a long cylinder such as a ground rod or an antenna.

The resistance between metallic conductors buried, or partially buried, in the ground, is capable of calculation if it is assumed that the ground has uniform resistivity. While this assumption is seldom realized in practice, the calculations are very useful in the design of ground conductors. The problem is identical with that of the calculation of the capacitance of the same conductors. The flow of current between the conductors follows the same laws as the flow of dielectric flux. It is more convenient in many capacitance problems to calculate the repulsion between the charges rather than to solve the flow problem, and that method will be used in the problems described in this paper.

Expressions are available for the capacitance of very few shapes of finite bodies. The capacitance of an isolated sphere, in absolute electrostatic units, or statfarads, is equal to the number of centimeters in its radius. The capacitance of two spheres near together is known, but is obtained by a rather complicated process. See "Theory of Alternating Currents" by Alex. Russell,

* Massachusetts Institute of Technology, and New England Power Construction Company.

ed. of 1914, Chapter VIII. The capacitance of two or more spheres far apart is more readily computed, as may be noted from examples in this paper.

A formula is available for the capacitance of an ellipsoid of revolution. It is given in Scientific Paper No. 568 of the Bureau of Standards, by F. W. Grover, page 576. (See also eq. 17 of this paper.)

The capacitance of a thin round plate is [1]

$$C = \frac{2a}{\pi}, \tag{1}$$

where a is the radius.

CYLINDER

A case occurring very frequently is that of a cylinder, the same general problem being encountered when a rod is driven vertically into the ground and when a rod or wire is buried in a horizontal position.

Two approximate formulas have been offered for this problem, which give results differing by 5 per cent for a ¾ inch, 8 foot ground rod, and by 3 per cent for an antenna wire ½ cm. in diameter and 100 meters long. See equations (17) and (19) of this paper and the references given with them.

In this paper a calculation for the capacitance of a rod of finite length is given, which is in the form of convergent series. It gives precise results when large numbers of terms are computed. The calculation follows the general method used by the writer in calculating the capacitance of a long, three-conductor cable.[2] Uniform distribution of charge is first assumed and then other distributions of charge are successively added, so as to keep the potential of the conductor the same at all points.

In Figure 1 is indicated a circular cylinder of length $2L$ cm. and radius a cm.

[1] "The Newtonian Potential Function," by B. O. Peirce, ed. of 1902, page 161.

[2] "The Direct Method of Calculation of Capacitance of Conductors," by H. B. Dwight, Trans. A. I. E. E., 1924, page 958.

First find the potential at the point P on the cylinder due to uniform charge density q per cm. along the cylinder. The potential at P due to the ring dy is [3]

$$\frac{qdy}{a}\left[\frac{a}{r}P_0\left(\frac{y}{r}\right) - \frac{1}{2}\frac{a^3}{r^3}P_2\left(\frac{y}{r}\right) + \frac{1.3}{2.4}\frac{a^5}{r^5}P_4\left(\frac{y}{r}\right) - \cdots\right], \quad (2)$$

where

$$P_0\left(\frac{y}{r}\right) = 1,$$

$P_2, P_4 \ldots$ represent zonal harmonics whose values are all less than 1 and where $r^2 = a^2 + y^2$.

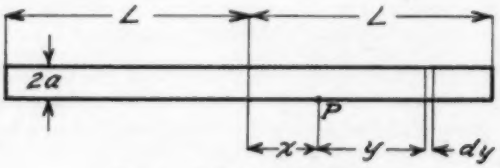


FIG. 1.—Cylindrical Conductor.

Where cylinders are dealt with in this paper, the length is about 100 or more times the radius, and all except the first term in a power series in a such as (2) may be neglected. Integrating the first term from $y = 0$ to $L - x$, the potential at P due to the part of the rod to the right is

$$q \operatorname{logn} \frac{L - x + \sqrt{a^2 + (L - x)^2}}{a}.$$

where logn denotes natural logarithm.

This becomes, by expanding and omitting terms in a except logn a,

$$q \operatorname{logn} \frac{2L}{a} - q\left(\frac{x}{L} + \frac{x^2}{2L^2} + \frac{x^3}{3L^3} + \frac{x^4}{4L^4} + \cdots\right).$$

The potential at P due to the part of the rod to the left is

[3] "Fourier Series and Spherical Harmonics," by W. E. Byerly, page 153, eq. (6).

obtained by changing x to $-x$. The total potential at P due to the uniform charge q per cm is

$$2q \log \frac{2L}{a} - q\left(\frac{x^2}{L^2} + \frac{x^4}{2L^4} + \frac{x^6}{3L^6} + \cdots\right). \qquad (3)$$

It is now desirable to find the potential at P due to a charge whose density varies as the square of the distance from the mid-section of the rod. The charge on the circle dy will then be

$$q_1 \frac{(x+y)^2}{L^2} dy$$

and the potential at P due to it is

$$\frac{q_1}{L^2} \frac{(x+y)^2}{\sqrt{a^2+y^2}} dy,$$

omitting terms with a in the numerator.

Expanding the numerator, and integrating the three terms from $y = 0$ to $L - x$, we obtain

$$\frac{q_1}{L^2}\left[\frac{y}{2}\sqrt{a^2+y^2} - \frac{a^2}{2}\log\,(y + \sqrt{a^2+y^2}) + 2x\sqrt{a^2+y^2} + x^2 \log\,(y + \sqrt{a^2+y^2})\right]_{y=0}^{L-x}.$$

Expanding this and adding the corresponding expression for the potential at P due to the left hand part of the rod, we obtain: Potential at P due to charge $q_1\ u^2/L^2$

$$= q_1\left[1 + \frac{x^2}{L^2}\left(2 \log \frac{2L}{a} - 3\right) - \frac{x^4}{L^4} - \frac{x^6}{2L^6} - \frac{x^8}{3L^8} - \cdots\right]. \qquad (4)$$

Similarly, the potential at P due to charge $q_2\ u^4/L^4$

$$= q_2\left[\frac{1}{2} + \frac{x^2}{L^2} + \frac{x^4}{L^4}\left(2 \log \frac{2L}{a} - \frac{25}{6}\right) - \frac{x^6}{L^6} - \frac{x^8}{2L^8} - \frac{x^{10}}{3L^{10}} - \cdots\right] \qquad (5)$$

and the potential at P due to charge $q_3\ u^6/L^6$

$$= q_3\left[\frac{1}{3}+\frac{x^2}{2L^2}+\frac{x^4}{L^4}+\frac{x^6}{L^6}\left(2\log_n\frac{2L}{a}-\frac{49}{10}\right)\right.$$
$$\left.-\frac{x^8}{L^8}-\frac{x^{10}}{2L^{10}}-\frac{x^{12}}{3L^{12}}-\cdots\right]. \quad (6)$$

Applying these equations to a rod ¾ inch in diameter and 30 feet in length, so that $L = 180$ inches, the equations become:
Potential due to q per cm.

$$= q\left(13.734-\frac{x^2}{L^2}-\frac{x^4}{2L^4}-\frac{x^6}{3L^6}-\frac{x^8}{4L^8}-\cdots\right). \quad (7)$$

Potential due to $q_1\ u^2/L^2$

$$= q_1\left(1+10.734\frac{x^2}{L^2}-\frac{x^4}{L^4}-\frac{x^6}{2L^6}-\frac{x^8}{3L^8}-\cdots\right). \quad (8)$$

Potential due to $q_2\ u^4/L^4$

$$= q_2\left(\frac{1}{2}+\frac{x^2}{L^2}+9.567\frac{x^4}{L^4}-\frac{x^6}{L^6}-\frac{x^8}{2L^8}-\cdots\right). \quad (9)$$

Potential due to $q_3\ u^6/L^6$

$$= q_3\left(\frac{1}{3}+\frac{x^2}{2L^2}+\frac{x^4}{L^4}+8.834\frac{x^6}{L^6}-\frac{x^8}{L^8}-\cdots\right). \quad (10)$$

Let $10.734q_1 = q$.
The potential due to

$$\frac{q}{10.734}\frac{u^2}{L^2}$$

is

$$q\left(.093+\frac{x^2}{L^2}-.093\frac{x^4}{L^4}-.047\frac{x^6}{L^6}-.031\frac{x^8}{L^8}-\cdots\right). \quad (11)$$

Adding (7) and (11), the potential due to the first 2 charges is

$$q\left(13.827-.593\frac{x^4}{L^4}-.380\frac{x^6}{L^6}-.281\frac{x^8}{L^8}-\cdots\right). \quad (12)$$

Let $9.567\, q_2 = .593q$,

$$q_2 = .062q.$$

Potential at x due to $.062q$ u^4/L^4

$$= q\left(.031 + .062\frac{x^2}{L^2} + .593\frac{x^4}{L^4} - .062\frac{x^6}{L^6} - .031\frac{x^8}{L^8} - \cdots\right). \quad (13)$$

Adding (12) and (13), the potential due to the first 3 charges is

$$q\left(13.858 + .062\frac{x^2}{L^2} - .442\frac{x^6}{L^6} - .312\frac{x^8}{L^8} - \cdots\right). \quad (14)$$

Similarly, the potential due to the first 4 charges is

$$q\left(13.875 + .087\frac{x^2}{L^2} + .050\frac{x^4}{L^4} - .362\frac{x^8}{L^8} - \cdots\right). \quad (15)$$

Applying this process again to the terms in x^2/L^2, x^4/L^4 and x^8/L^8, the potential due to the first 7 charges is

$$q\left(13.863 - .007\frac{x^2}{L^2} - .001\frac{x^4}{L^4} - .355\frac{x^8}{L^8} - \cdots\right). \quad (16)$$

The average potential, by integrating from 0 to L, is

$$q(13.863 - .002 - 0 - .040 - \cdots) = 13.82q.$$

The total charge on the rod is obtained by integrating from 0 to L and multiplying by 2 and is

$$(1 + .031 + .012 + .007 - .003 - .001 - 0 - \cdots) \times 360 \times 2.54q = 1.046 \times 360 \times 2.54q,$$

$$\frac{1}{C} = \frac{13.82}{1.046 \times 360 \times 2.54} = \frac{13.21}{360 \times 2.54}.$$

This result can be compared with two published formulas for the capacitance of a cylinder. One approximation is to consider the cylinder to be nearly the same as an ellipsoid of revolution and to use the formula for that shape for the case when the diam-

eter is small compared with the length. This is [4] in the notation of this paper

$$\frac{1}{C} = \frac{1}{L} \operatorname{logn} \frac{2L}{a}, \tag{17}$$

which gives, for the problem just considered,

$$\frac{1}{C} = \frac{13.73}{360 \times 2.54}.$$

The difference is 3.9 per cent.

Another approximate calculation is to assume uniform charge density and to calculate the average potential. If terms in a, except the term logn a, are omitted from eq. (2), as was previously done in this paper, there is obtained [5]

$$\frac{1}{C} = \frac{1}{L}\left(\operatorname{logn} \frac{2L + \sqrt{a^2 + 4L^2}}{a} + \frac{a}{2L} - \sqrt{1 + \frac{a^2}{4L^2}} \right). \tag{18}$$

Terms in a, except logn a, should not be left in the formula if other terms in a have been neglected in deriving it.

Eq. (18) therefore becomes

$$\frac{1}{C} = \frac{1}{L}\left(\operatorname{logn} \frac{4L}{a} - 1 \right) \tag{19}$$

This gives, for the numerical problem just considered,

$$\frac{1}{C} = \frac{13.12}{360 \times 2.54}.$$

This is different from the value calculated in this paper by less than 1 per cent. Formula (19) is preferable to (17) since it is more accurate and no more difficult to use. Eq. (19) is therefore used in this paper for the calculation of ground resistance.

The potential at the center of the rod of length $2L$, due to the

[4] See Technologic Paper No. 108 of the Bureau of Standards, "Ground Connections for Electrical Systems," 1918, page 221.

[5] "On the Capacity of Radio-Telegraphic Antennae," by G. W. O. Howe, The Electrician, Vol. 73, 1914, pages 829, 859 and 906.

uniform charge q per cm., is found by putting $x = 0$ in eq. (3) and is

$$2q \log_n \frac{2L}{a},$$

which gives $13.73q$ for the case considered. The average potential, from eq. (19), is

$$\frac{2qL}{C} = 2q\left(\log_n \frac{4L}{a} - 1\right),$$

which gives $13.12q$ for the same numerical case. The potential at the center of the rod is therefore 4.7 per cent higher than the average potential due to uniform charge density.

As a second numerical example, an antenna wire of radius .25 cm. and length 10,000 cm. (L = 5,000 cm.), may be taken. By the successive approximation calculation of this paper, $1/C$ = .00207. By eq. (17), $1/C$ = .00212, which is 2.5 per cent larger, and by eq. (19) $1/C$ = .00206, which is 0.5 per cent smaller.

The successive approximation method described in this paper can be used with groups of a few cylinders.

It would not be expected that the calculation of the average potential due to uniform charge distribution as in (19), would necessarily give a close approximation to the capacitance of a rod, for the potential at the end of the rod is nearly 50 per cent less than the potential at the middle, when uniform charge is assumed. The successive approximation calculation is needed to justify eq. (19). In other cases discussed later in this paper, the highest and lowest potentials do not differ by more than a few per cent, uniform charge distribution being assumed. In such cases the average potential method can be expected, and is shown, to give values of capacitance which are very close approximations to the true values. On the other hand, in the case of a thin, round plate the average potential method gives a value of C about 8 per cent too low.

A successive approximation method using mechanical integra-

tion has been used by Dr. F. W. Grover in Scientific Paper of the Bureau of Standards, No. 568, 1928, page 577, to show that the average potential method is correct to within 2 or 3 parts in 1000 for a long straight antenna wire.

Since this paper was written, Dr. Grover has communicated to me the following series which he has adapted from eq. 20 (b) of the paper "Lösung zweier Potentialprobleme der Elektrostatik," by E. Hallen, Arkiv för Matematik, Astronomi och Fysik, Vol. 21A, No. 22, 1929, Stockholm:

$$\frac{C}{L} = \frac{2}{\omega} + \frac{1.22741}{\omega^2} + \frac{2.17353}{\omega^3} + \frac{11.0360}{\omega^4} \cdots, \qquad (19A)$$

where

$$\omega = 2 \log_n \frac{2L}{a}.$$

This gives, for the first example,

$$\frac{1}{C} = \frac{13.02}{360 \times 2.54}.$$

to compare with

$$\frac{1}{C} = \frac{13.12}{360 \times 2.54}.$$

by the average potential method. For the second example, it gives

$$\frac{1}{C} = .00205.$$

to compare with .00206 by the average potential method. It is to be expected that the average potential method would give too large a value for $1/C$ since it does so in the case of a thin disk (see the sentence following eq. 31).

The resistance to earth of a conductor deeply embedded in the earth is [6]

$$R = \frac{\rho}{4\pi C}, \qquad (20)$$

[6] Technologic Paper No. 108 of the Bureau of Standards, "Ground Connections for Electrical Systems," Appen. III, p. 219, and "An Investigation of Earthing Resistances," by P. J. Higgs, Jour. I. E. E., England, Vol. 68, 1930, Appendix, p. 745.

where C is the capacitance in electrostatic units, or statfarads, of the conductor when isolated in air, and ρ is the resistance per cm. cube of the earth. The same expression gives the resistance between two bodies deeply buried in the earth, if C is the capacitance of the two bodies in air, when one is positive and one negative.

If the bodies are near the surface of the ground, the flow lines at the plane of the ground surface lie in that plane, and the flow problem corresponds to the capacitance of the body or bodies together with their images in the ground surface. If C includes the capacitance of the images, the resistance of the body or bodies is given by

$$R = \frac{\rho}{2\pi C}. \tag{21}$$

The resistance to the earth of a rod driven vertically to a depth L cm. in the earth is given by (19) and (21) and is

$$R = \frac{\rho}{2\pi L}\left(\operatorname{logn}\frac{4L}{a} - 1\right). \tag{22}$$

The resistance of a buried straight rod or wire, no part of which is near the surface (usually a horizontal conductor) is given by (19) and (20) and is

$$R = \frac{\rho}{4\pi L}\left(\operatorname{logn}\frac{4L}{a} - 1\right), \tag{23}$$

where the length of the rod or wire is $2L$ cm.

Two Ground Rods in Parallel

As vertical ground rods are used to a considerable extent, because of their ability to reach permanent moisture, calculations for ground rods in parallel will be described.

Let two parallel vertical rods of depth L cm. and radius a cm. be separated by s cm. between centers and be connected electrically in parallel. Then it is required to find the capacitance of these rods and their images, that is, of two rods of length $2L$. Assume that both rods have uniform charge q per cm.

The potential at distance x from the center of one rod due to the charge qdy on the other rod is

$$\frac{qdy}{\sqrt{s^2 + y^2}}.$$

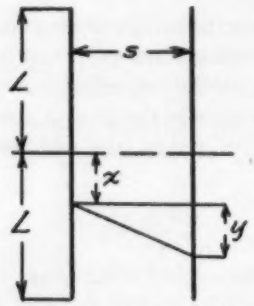


FIG. 2.—Two Rods Connected in Parallel.

The potential at x due to uniform charge q per cm. on the other rod is obtained by integrating from $y = 0$ to $L + x$ and from $y = 0$ to $L - x$, and is

$$q \log_n \frac{L + x + \sqrt{s^2 + (L + x)^2}}{s} + q \log_n \frac{L - x + \sqrt{s^2 + (L - x)^2}}{s}. \quad (24)$$

The average potential on one rod due to uniform charge on the other is obtained by multiplying by dx/L and integrating from $x = 0$ to L, and is

$$q\left[2 \log_n \frac{2L + \sqrt{s^2 + 4L^2}}{s} + \frac{s}{L} - \frac{\sqrt{s^2 + 4L^2}}{L} \right]. \quad (25)$$

Adding the average potential of the rod due to its own charge, and dividing by $4qL$, the value of $1/C$ for the pair of rods is given.

Then, by (21), the resistance to earth of the pair of rods is

$$R = \frac{\rho}{4\pi L}\left[\log_n \frac{4L}{a} - 1 + \log_n \frac{2L + \sqrt{s^2 + 4L^2}}{s} + \frac{s}{2L} - \frac{\sqrt{s^2 + 4L^2}}{2L}\right]. \quad (26)$$

A short alternative calculation, which is good for spacings of 20 feet or more, is to replace each ground rod by a half-buried sphere which is equivalent in resistance to one isolated ground rod. The resistance to earth of a ground rod of ¾ inch diameter and 10 feet depth is, by (22),

$$R = \frac{\rho}{2\pi L} \times 6.155.$$

The hemisphere which is buried and its image above the ground surface, make up a complete sphere whose capacitance as an isolated conductor is equal to its radius A in cm. By (21), the resistance of the hemisphere to earth is

$$R = \frac{\rho}{2\pi A}.$$

Therefore, the radius of the hemisphere which is equivalent to the isolated ground rod is

$$A = \frac{L}{6.155} = 1.625 \text{ ft.}$$

The capacitance of two spheres at a distance s center to center, connected in parallel, is quickly calculated when s is not small. Let the spheres carry equal charges q.

The potential of the surface of each sphere is

$$q\left(\frac{1}{A} + \frac{1}{s}\right),$$

where the dimensions are in centimeters. If the distance between

the centers is 10 feet,

$$\frac{1}{C} = \frac{1}{2 \times 12 \times 2.54}\left(\frac{1}{1.625} + \frac{1}{10}\right) = \frac{.715}{2 \times 12 \times 2.54}.$$

For two spheres widely separated

$$\frac{1}{C} = \frac{1}{2 \times 12 \times 2.54}\left(\frac{1}{1.625}\right) = \frac{.615}{2 \times 12 \times 2.54}.$$

The ratio of the conductivity to earth of the two spheres 10 feet apart to that of two isolated spheres is

$$\frac{.615}{.715} = .86.$$

The ratio calculated by (26) is .88, showing 2.5 per cent difference at 10 feet spacing. The difference is .5 per cent for 20 feet spacing. The conclusion may therefore be stated that for two 10 foot ground rods the short, sphere method of calculation is a good approximation for spacings of 20 feet or more. With three or four 10 ft. ground rods, with 10 ft. spacing at the corners of an equilateral triangle or a square, the error due to the sphere method is approximately 4 per cent, and less with larger spacings.

Groups of Ground Rods

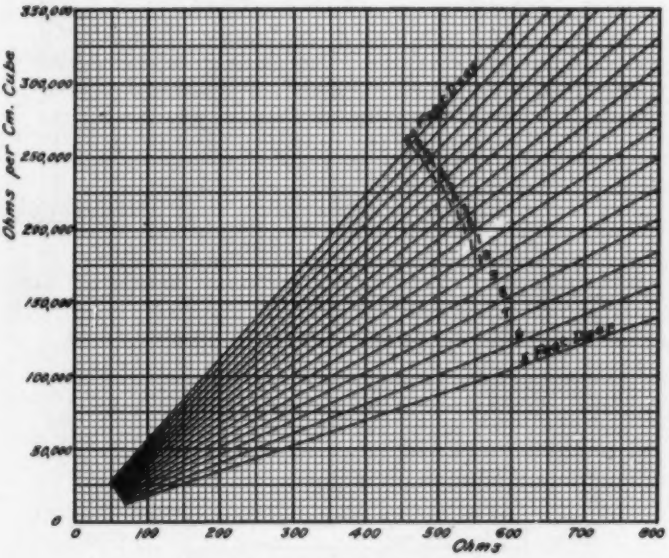
The potential of a rod due to its own charge and the charges of several other rods can be found by several applications of (25). Similarly, the potential of the surface of a sphere due to its own charge and the charges of several other spheres is

$$\frac{q_1}{A} + \frac{q_2}{s_2} + \frac{q_3}{s_3} + \cdots,$$

where for an approximate calculation q_1, q_2, q_3, etc., may be assumed equal.

Ground rods are used singly and in small groups for measurement purposes, and in large and small groups for permanent grounds as at substations. It is desirable to be able to compare

various groups of rods, assuming uniform conductivity of the soil. For this purpose, sets of curves are given in Figs. 3 to 6. From them it is possible to estimate how many ground rods and how much area will be required for a certain number of ohms to ground in a given locality, from measurements on a few temporary test ground rods.



For low resistances, divide both scales by 10.
For high resistances, multiply both scales by 10.

FIG. 3.—Resistance of One Ground Rod of ¾ Inch Diameter.

The basis of comparison used is the resistance of one isolated ground rod of ¾ inch diameter and 10 ft. depth. By means of Fig. 3, the resistance of single rods of various depths can be found, and, if desired, the resistivity of the soil. In Fig. 4 the conductivity of 2, 3 and 4 rods is given in terms of the conductivity of isolated rods, and in Fig. 5 larger numbers of rods are covered.

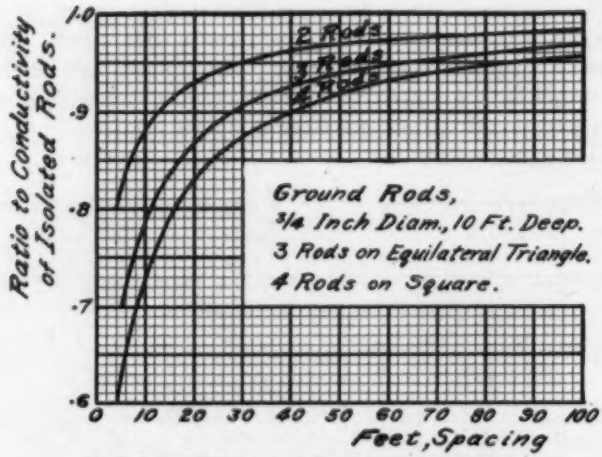


FIG. 4.—Conductivity of Ground Rods in Parallel.

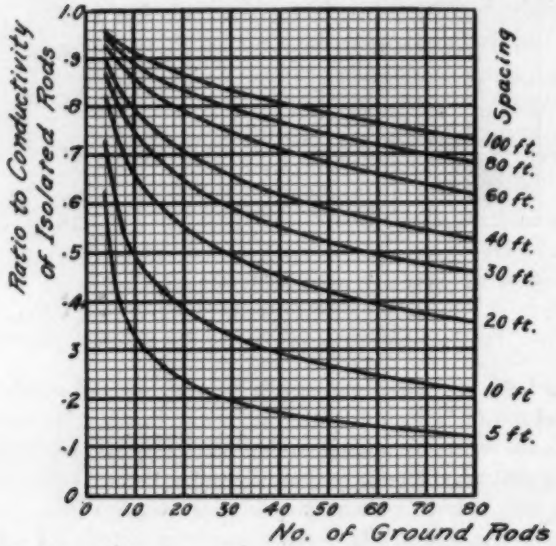


FIG. 5.—Conductivity of Ground Rods in Parallel, on an Area.

Figure 6 gives the information in different form, and shows the lowest resistance that can be obtained from a given area. This information is often of value, as it may save one from the wasteful attempt of putting additional ground rods in an area which could not give the desired low resistance even if an infinite number of

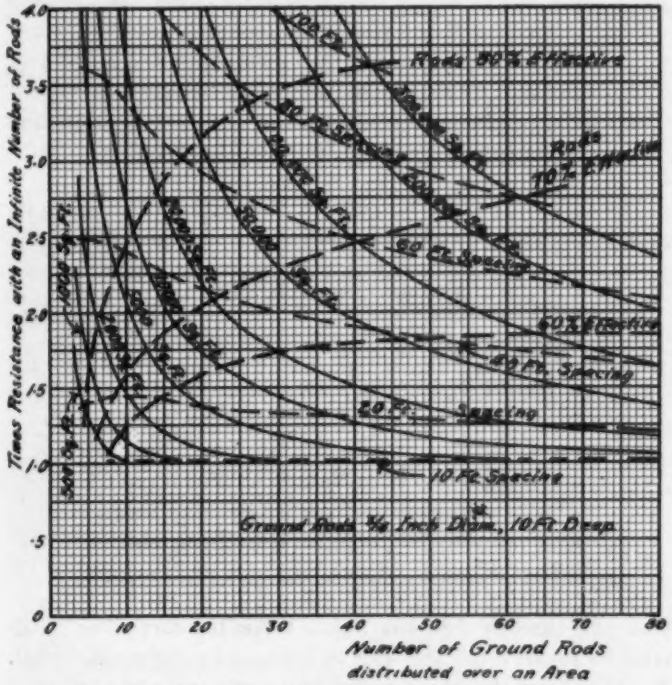


FIG. 6.—Conductivity of Ground Rods in Parallel, on an Area.

rods were used. For example, if 25 ground rods are distributed over a square area of 5,000 sq. ft., Fig. 6 shows that the resistance is 1.25 times the resistance of an infinite number of rods in the same area. No matter what the resistance in ohms may be, if it is desired to have less than 80 per cent of the resistance with 25

rods there is no use in putting more rods in the area of 5,000 sq. ft., but it is necessary to utilize a larger area.

Fig. 6 shows also that if there are more than 10 rods, 10 feet deep, there is no use in using closer than 10 ft. spacing.

It is usually desirable not only to use the ground area effectively, but to utilize the rods effectively, since they involve considerable cost. If one wishes the rods to be at least 70 per cent as effective as they would be if isolated, that is, to have at least 70 per cent of the conductivity of isolated rods, it is seen from Figs. 5 and 6 that it is necessary to use over 30 ft. spacing for 16 rods in a square area, over 40 ft. for 25 rods and over 60 ft., for 49 rods.

The curves are based on uniform distribution of ground rods over square areas, as indicated in Fig. 7. The boundary of the

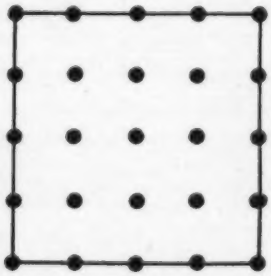


FIG. 7.—Distribution Used for Calculating the Curves.

area runs through the outer rods. While the curves are calculated for square areas, they give an estimate for rectangular areas. A reading from the curves applies to a single area which contains all the ground rods in the vicinity.

In the calculation of cases involving more than four spheres, equal charges were assumed for the spheres. The potential of a corner sphere was calculated, and that of an innermost sphere. The average of the two was taken as the average potential of the group. This procedure was checked by finding the actual charges

on a small number of spheres by simultaneous equations. The capacitance found in this way checked closely with the result of the approximate calculation described. For a large number of spheres, the results of Figs. 5 and 6 are approximate.

The effect of the buried wires used to connect the ground rods together has not been included in the calculations. They usually lie nearer the surface of the ground, and therefore in earth of higher resistivity, than the vertical ground rods. If the conductivity of each ground rod be assumed to be increased by the same percentage because of the connecting wires, then the wires will have very little effect on the comparison of different groups of rods as shown by the curves.

Metallic Strip

The resistance to earth of a buried, horizontal, long, straight metallic strip of rectangular cross section can be expressed by formulas similar to (19) and (20) which apply to buried, horizontal wires. For deriving the formulas, the method of finding the average potential due to uniform charge distribution will be used.

Let the length of the strip be $2L$ cm., its width be a cm., and its thickness b cm., let a and b be small compared with L and let b be small compared with a, that is, less than about one-eighth a. Let the uniform charge density be q per sq. cm.

The average potential of face a due to its own charge is

$$q\left(2a \operatorname{logn}\frac{4L}{a} + a\right)$$

due to the charge on the two faces of width b

$$q\left(4b \operatorname{logn}\frac{4L}{a}\right)$$

and due to the charge on the opposite face of width a

$$q\left(2a \operatorname{logn}\frac{4L}{a} + a - 2\pi b\right).$$

The average potential of face a is, therefore,

$$q\left\{4(a+b)\operatorname{logn}\frac{4L}{a}+2a-2\pi b\right\}.$$

The average potential of face b due to its own charge is

$$q\left(2b\operatorname{logn}\frac{4L}{b}+b\right)$$

due to the two faces of width a

$$q\left(4a\operatorname{logn}\frac{4L}{a}-\pi b\right)$$

and due to the opposite face of width b,

$$q\left(2b\operatorname{logn}\frac{4L}{a}-2b\right).$$

The average potential of face b is,

$$q\left\{4(a+b)\operatorname{logn}\frac{4L}{a}-2b\operatorname{logn}\frac{b}{a}-(\pi+1)b\right\}.$$

The average potential of the strip is, neglecting terms in b^2,

$$q\left[4(a+b)\operatorname{logn}\frac{4L}{a}+\frac{2a^2-2\pi ab}{a+b}\right].$$

The total charge on the strip is $4qL(a+b)$.

$$\frac{1}{C}=\frac{1}{L}\left[\operatorname{logn}\frac{4L}{a}+\frac{a^2-\pi ab}{2(a+b)^2}\right]. \tag{27}$$

The first term of the bracket of (27) is in agreement with the formula for a deeply buried strip, in the article by R. Rüdenberg, Elektrotechnische Zeitschrift, Vol. 46, Sept. 3, 1925, page 1342. This article gives also a formula for a buried ring of wire.

If it is desired to change the values of capacitance and inductance of an antenna in some other way than by using round con-

ductors of different sizes, strip conductor might be used and the capacitance calculated by (27).

Example: Length of strip $= 2L = 60$ ft.,
width $= a = 1.5$ inches,
thickness $= b = .25$ inches.

$$\frac{1}{C} = \frac{6.87 + .17}{30 \times 12 \times 2.54} = .0077.$$

$$R = \frac{\rho}{4\pi C}$$

for a completely buried strip.

Rectangular and Round Plates

The capacitance of a thin rectangular plate a cm. $\times$ b cm., whose width is not small compared with its length, as obtained by the average potential method, is given by [7]

$$\frac{1}{C} = 2\left[\frac{1}{a}\log_n \frac{a + \sqrt{a^2 + b^2}}{b} + \frac{1}{b}\log_n \frac{b + \sqrt{a^2 + b^2}}{a} + \frac{a}{3b^2} + \frac{b}{3a^2} - \frac{(a^2 + b^2)\sqrt{a^2 + b^2}}{3a^2b^2}\right]. \quad (28)$$

The case of a rectangular or square plate is not quite so favorable for the application of the average potential method as the case of a long, narrow conductor. However, since the discrepancy is not more than about 8 per cent, the calculation is good for most practical purposes.

The relative accuracy of (28) for a square or nearly square plate can be judged by investigating the case of a thin, round plate. Using the formulas for potential due to a ring given in "Fourier's Series and Spherical Harmonics," by W. E. Byerly, ed. of 1893, pages 11 and 153, the average potential of a thin round plate of radius a cm., due to uniform charge density q per

[7] "Capacity of Rectangular Plates," by G. W. O. Howe, Radio Review, 1920, Vol. 1, page 710.

sq. cm., is given by three series which add up to the following:

$$V(\text{aver.}) = \frac{2\pi qa}{3}\left(2 + \frac{1}{4} + \frac{3^2}{4^2.6} + \frac{3^2.5^2}{4^2.6^2.8} + \frac{3^2.5^2.7^2}{4^2.6^2.8^2.10} + \cdots\right). \quad (29)$$

This series is convergent, but rather slowly so. Its sum can be found by comparing it with the series

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \frac{1}{7^2} + \cdots. \quad (30)$$

Multiplying the bracket of (29) by 0.4094 to make the 6th terms alike, it is found that $1/C$ or $V(\text{aver.})/\pi a^2 q$ is very slightly greater than

$$\frac{1.696}{a}.$$

Multiplying the bracket of (29) by 0.2931 to make its 6th term equal to the 7th term of (30) and all its succeeding terms distinctly less than the corresponding terms of (30), it is found that $1/C$ is less than $1.710/a$. The difference between these two limiting values is 0.8 per cent, and the value of $1/C$ as given by the average potential method is very nearly equal to the smaller limit, namely

$$\frac{1.696}{a}. \quad (31)$$

The value of $1/C$ is, however,

$$\frac{\pi}{2a} = \frac{1.571}{a},$$

by eq. (1). The average potential method, therefore, gives a value of $1/C$ for a thin round plate, which is 8 per cent too high. Dr. F. W. Grover has recently communicated to me that in 1925 he obtained the following result for a round disk carrying uniform

charge density, by using mechanical integration,

$$\frac{V(\text{aver.})}{\pi a^2 q} = \frac{1.6966}{a},$$

which confirms the above result.

It is of interest to compute the potential of the center of the plate, due to uniform charge distribution. It is $2\pi aq$. The average value is, as found above,

$$\frac{1.696}{a} \times \pi a^2 q = 1.696\pi aq.$$

The potential of the center is 18 per cent higher than the average potential. This comparatively large percentage may account to some extent for the 8 per cent error in this case due to the use of the average potential method. It may be remembered that in the case of the rod calculated in this paper the potential of the middle point of the rod was only 4.7 per cent higher than the average potential due to uniform charge distribution, and the average potential method gave very accurate results.

The potential of the center of a thin rectangle, a cm by b cm., is

$$2q\left(a \log_n \frac{b + \sqrt{a^2 + b^2}}{a} + b \log_n \frac{a + \sqrt{a^2 + b^2}}{b}\right). \quad (32)$$

The potential of a corner of the rectangle is ½ as much.

The average potential of a square plate of side b cm. is, by (28), $2.973\,bq$. The potential of the center, by (32), is $3.525bq$, which is 18.6 per cent greater than the average potential. If the square of side b cm. has the same area as a circle of radius r cm., the average potential of the square plate is $2.973qr\sqrt{\pi}$ and

$$\frac{1}{C} = \frac{1.677}{r}.$$

This is 1.1 per cent less than $1.696/r$ for the circular plate.

All the above figures for a square plate correspond so closely to

those for a circular plate of the same area that the conclusion may be drawn that the true value of $1/C$ for a square plate is approximately 8 per cent less than the value given by (28) by the average potential method.

When the length of a rectangular plate is 5 times its width, the potential at the center is found from equations (28) and (32) to be 15 per cent higher than the average potential, thus indicating that the error due to the average potential method is nearly 8 per cent. When the length is 10 times the width, the difference of the two potentials is 13 per cent, indicating almost as large an error as before.

When the length is 100 times the width, the difference of the two potentials is 8.6 per cent. The average potential method seems more suitable for long, narrow rectangles and for long cylinders than for squares.

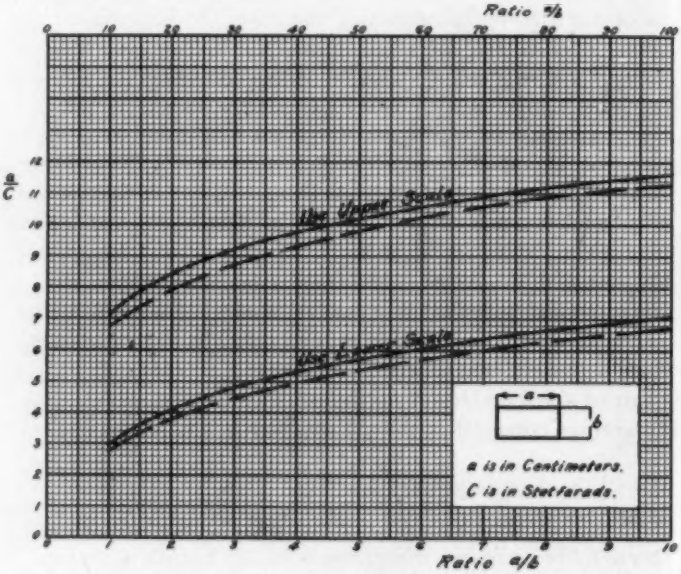


FIG. 8.—Capacitance of a Rectangular Plate.

The capacitance of rectangles as calculated by the average potential method can be obtained from the full lines of Fig. 8. These values of a/C are too large, and the broken lines show an estimated correction.

In conclusion, the average potential method of calculating capacitance and resistance to ground is a very useful device. It is most accurate with a long, straight conductor, and gives errors up to about 8 per cent with a thin, rectangular or round plate.

Resistance to Alternating Current

The resistance to alternating current of a circuit consisting of an overhead wire several miles long with the earth as a return path, is a special problem, involving proximity effect. Formulas for this problem are given by J. R. Carson in the Bell System Technical Journal, October, 1926. By putting the frequency equal to 60 cycles in his eq. (34), and noting that r of his formula is small for usual values of ground resistivity, it is found that the resistance of the path through the earth is 1/10 ohm per mile, for 60-cycle current, and that this value is constant within a very few per cent, for high and low values of earth resistivity which may be encountered. The value is proportional to the frequency, at low frequencies. The greater part of the alternating current is in a path one or two miles wide and deep, parallel to the overhead wire. If the resistivity is high, the current spreads out and if it is low, the current flows closer to the overhead wire. The calculation is based on uniform resistivity and it is probable that extreme local changes in resistivity would affect the above result, but that gradual changes would make very little difference in it. If there is an overhead ground wire or cable, its impedance is in parallel with that of the earth path.

The calculations previously described in this paper have been based on direct current, but as the earth paths considered are usually considerably shorter than one mile, differences between 60-cycle alternating current and direct current would scarcely be appreciable. The 60-cycle resistance of the earth path be-

tween two stations is the sum of the resistances to earth of the stations plus 1/10 ohm per mile for the distance between them, measured along the transmission line. It should be noted, however, that the resistance is difficult both to measure and to calculate for the common cases where there are overhead ground conductors on steel towers, unless the terminal resistances to ground are extremely low.

ON TAUBERIAN THEOREMS OF HARDY AND LITTLEWOOD AND A NOTE ON WINTNER'S PAPER

By Shikao Ikehara

Norbert Wiener [1] developed a general method for proving theorems called Tauberian by Hardy and Littlewood. Theorems are in the form: if a function has a mean value of a certain type for large values of its arguments, and if the rate of growth of the function is subject to certain restrictions, then the function has a limiting value at infinity. By means of theorems of this type N. Wiener proves the O_L Tauberian theorems for Abel, Lambert, and Borel summation. Moreover, he [2] shows an application to Fourier series. In case of the summability of Fourier series and integrals S. B. Littauer [3] is able to apply the general method successfully. In this paper we shall show the simplicity and powerfulness of the new machinery applying to certain problems.

The general Tauberian theorem by N. Wiener reads:

Key Theorem. *Let $K_1(x)$ and $K_2(x)$ be two bounded functions which are $O\left(\frac{1}{x(\log x)^2}\right)$ at 0 and ∞. Let $K_1(x)$ be measurable and non-negative, and let*

$$\int_0^\infty K_1(x)x^{iu}dx \neq 0 \quad [-\infty < u < \infty].$$

Let $K_2(x)$ be continuous, except for a finite number of finite jumps. Let $g(x)$ be a measurable function bounded below and such

[1] N. Wiener, "A New Method in Tauberian Theorems," Jour. of Math. and Phys., Mass. Inst. of Tech., 7 (1928), pp. 161–184.

[2] N. Wiener, "A Type of Tauberian Theorem Applying to Fourier Series," *Proc. London Math. Soc.* (2), 30 (1929), pp. 1–8.

[3] S. B. Littauer, "A New Tauberian Theorem with Application to the Summability of Fourier Series and Integrals," Jour. of Math. and Phys., Mass. Inst. of Tech., 8 (1929), pp. 216–234.

PHYSICS LIBRARY

that

$$(a) \qquad \lim_{\lambda \to 0[\infty]} \int_0^\infty g(\lambda x) K_1(x)dx = W \int_0^\infty K_1(x)dx,$$

$$(b) \qquad \int_0^\infty g(\lambda x) K_1(x)dx$$

is bounded. $[0 < \lambda < \infty]$. *Then*

$$(c) \qquad \lim_{\lambda \to 0[\infty]} \int_0^\infty g(\lambda x) K_2(x)dx = W \int_0^\infty K_2(x)dx.$$

$[\infty]$ indicates that at these points it may be consistently substituted for 0.

I am deeply indebted to Professor Wiener for his suggestion of the topic and for his guidance in this paper.

§ 1. On Tauberian Theorems of Hardy and Littlewood

Recently Hardy and Littlewood [4] proved:

THEOREM 1. *If $f(t)$ is positive [5] and integrable [6] over every finite range $(0, T)$ and $e^{-xt}f(t)$ is integrable over $(0, \infty)$ for every $x > 0$; and if*

$$(1.1) \qquad g(x) = \int_0^\infty e^{-xt} f(t)dt \sim Hx^{-\alpha},$$

where $\alpha > 0$, $H > 0$, when $x \to 0$; then

$$(1.2) \qquad F(t) = \int_0^t f(t)dt \sim \frac{H}{\Gamma(1+d)} x^\alpha$$

when $x \to \infty$. The result is still true when $\alpha > 0$, $H = 0$ if we interpret (1.1) to mean $g(x) = o(x^{-d})$ and (1.2) similarly; and it is still true when $\alpha = 0$, $H > 0$.

We write (1.1)

$$\int_0^\infty x^\alpha e^{-xt} f(t)dt = \int_0^\infty x^{\alpha-1} e^{-xt} f(t)d(xt),$$

[4] G. H. Hardy and J. E. Littlewood, "Notes on the Theory of Series (XI): On Tauberian Theorems," *Proc. London Math. Soc.* (2), 30 (1929), pp. 23–37.

[5] In the wider sense.

[6] In the sense of Lebesgue.

and substituting $xt = \xi$ and $t^{1-\alpha}f(t) = g(t)$, we get

$$(1.11)\qquad \int_0^\infty \xi^{\alpha-1}e^{-\xi}g\left(\frac{\xi}{x}\right)d\xi = W\int_0^\infty \xi^{\alpha-1}e^{-\xi}d\xi = W\int_0^\infty K_1(\xi)d\xi.$$

Likewise, (1.2) can be transformed to

$$(1.21)\qquad \int_0^x \xi^{-\alpha-1}g\left(\frac{\xi}{x}\right)d\xi = W\int_0^x \xi^{-\alpha-1}d\xi = W\int_0^x K_2(\xi)d\xi.$$

Then we state

THEOREM 1_w. *Let K_1, K_2 and g satisfy the conditions laid down for them in the Key Theorem. If* (1.11) *is true, and*

$$\int_0^\infty \xi^{\alpha-1}e^{-\xi}g\left(\frac{\xi}{x}\right)d\xi$$

is bounded. $[0 < x < \infty]$ *Then* (1.21) *is true.*

THEOREM 2. *If $f(t)$ is positive and integrable over $(\delta, 1)$, for every $\delta > 0$, and $t^n f(t)$ is integrable over $(0, 1)$ for sufficiently large n; and if*

$$(1.3)\qquad g_n = \int_0^1 t^n f(t)dt \sim Hn^{-\alpha},$$

where $\alpha > 0$, $H > 0$, when $n \to \infty$ by integral values; then

$$(1.4)\qquad F(x) = \int_{1-x}^1 f(t)dt \sim \frac{H}{\Gamma(1+\alpha)}x^\alpha$$

when $x \to 0$. The result is still true when $\alpha > 0$, $H = 0$, if (1.3) *and* (1.4) *be then interpreted as in Theorem* 1.

Since [7] g_n decreases when n increases, we may replace $t = e^{-\omega}$, $e^{-\omega}f(e^{-\omega}) = h(\omega)$. We have

$$(1.31)\qquad g(x) = \int_0^\infty e^{-x\omega}h(\omega)d\omega \sim Hx^{-\alpha},$$

$$(1.41)\qquad H(x) = \int_0^x h(\omega)d\omega \sim \frac{H}{\Gamma(1+\alpha)}x^\alpha.$$

[7] Loc. cit., p. 24.

It is easy to see that Theorem 2 is intimately connected with Theorem 1_w, and in turn, it is equivalent to

THEOREM 3. *If $f(t)$ is positive and integrable over every finite range $(0, T)$, and $e^{-xt}f(t)$ is integrable over $(0, \infty)$ for sufficiently large x; and if* (1.1) *holds, with $\alpha > 0$, when $x \to \infty$; then* (1.2) *holds when $x \to 0$.*

This theorem differs from theorem 1 only in an exchange of roles between large and small values of x; the truth of the theorem is asserted in the Key Theorem.

THEOREM 4. *Suppose that $f(t)$ is positive and integrable over every finite range $(0, T)$, and that*

$$\frac{f(t)}{(t+x)^\rho},$$

where $\rho > 0$, is integrable over $(0, \infty)$ for some (and so for all) $x > 0$. Suppose further that

$$h(x) = \int_0^\infty \frac{f(t)dt}{(t+x)^\rho} \sim \frac{H}{x^\sigma}, \tag{2.1}$$

where $0 < \sigma < \rho$, $H > 0$, when $x \to \infty$. Then

$$F(x) = \int_0^x f(t)dt \sim \frac{H\Gamma(\rho)}{\Gamma(\sigma)\Gamma(\rho-\sigma+1)} x^{\rho-\sigma}, \tag{2.2}$$

when $x \to \infty$.

Furthermore, if (2.1) *holds when $x \to 0$, then* (2.2) *holds when $x \to 0$.*

Transforming (2.1) and (2.2) as before, we write

$$K_1(\xi) = \int_0^\infty \frac{1}{\xi^{\sigma-\rho+1}(1+\xi)^\rho} d\xi,$$

$$K_2(\xi) = \int_0^\infty \xi^{\sigma-\rho-1} d\xi,$$

$$g(t) = f(t)t^{\sigma-\rho+1}.$$

THEOREM 4_w. *Under the same conditions laid down for K_1, K_2 and g in the Key Theorem, if*

$$\lim_{x\to\infty}\int_0^\infty \frac{g(x\xi)}{\xi^{\sigma-\rho+1}(1+\xi)^\rho}d\xi = W\int_0^\infty \frac{1}{\xi^{\sigma-\rho+1}(1+\xi)^\rho}d\xi; \tag{2.11}$$

$$\int_0^\infty \frac{g(x\xi)}{\xi^{\sigma-\rho+1}(1+\xi)^\rho}d\xi \text{ is bounded } [0<x<\infty].$$

Then

$$\lim_{x\to\infty}\int_0^1 g(x\xi)\xi^{\sigma-\rho-1}d\xi = W\int_0^1 \xi^{\sigma-\rho-1}d\xi. \tag{2.21}$$

And we add that *the converse is also true.*

In order to indicate Wiener's method of proof, we prove Theorem 4_w.

$$K_1(\xi) = \int_0^\infty \frac{1}{\xi^{\sigma-\rho+1}(1+\xi)^\rho}d\xi = \frac{\Gamma(\sigma)\Gamma(\rho-\sigma)}{\Gamma(\rho)},$$

and

$$K_2(\xi) = \int_0^1 \xi^{\sigma-\rho-1}d\xi = \frac{1}{\sigma-\rho}.$$

Furthermore we know that

$$\int_0^\infty \frac{\xi^{i\delta}d\xi}{\xi^{\sigma-\rho+1}(1+\xi)^\rho} = \frac{\Gamma(\sigma-i\delta)\Gamma(\rho-\sigma+i\delta)}{\Gamma(\rho)}$$

has no zero, where δ is any real number. It is evident that

$$\int_0^\infty \frac{f(t)}{(t+x)}dt$$

is bounded for all $x>0$. Let us assume that

$$f(t) = 0 \qquad \text{for} \qquad 0 \le t \le 1,$$

but this restriction does not affect the asymptotic property of (2.1). When $x \to 0$, this expression is still bounded. Since $f(t)$ is positive, $g(t)$ is certainly non-negative. Thus they satisfy the Key Theorem.

Since [8]

$$F(x) \leq (2x)^{\rho} \int_0^x \frac{f(t)dt}{(t+x)^{\rho}} \leq (2x)^{\rho} \int_0^{\infty} \frac{f(t)dt}{(t+x)^{\rho}} = O(x^{\rho-\sigma}),$$

for large x, and

$$\int_0^1 \xi^{\sigma-\rho-1}\xi^{i\delta}d\xi \neq 0,$$

where δ is any real number, then the converse of Theorem 4_{ω} holds.

§ 2. A Note on Wintner's Paper

In the Laplacian problem of functions of large numbers we consider a moment integral of the type

$$\Xi(x) = \int_a^b (f(t))^x \Phi(t)dt \quad (-\infty < a \leqq t \leqq b < +\infty),$$

where $|f(t)|$ is constant in no part interval and which is dominated by large values of x. Laplace derived in several ways to which there are some objections a simple evaluation of $\Xi(x)$. Wintner [9] investigated the extent of the problem under rigorous conditions.

He calls the function $f(t)$ defined in $[0, \infty]$ unlimitedly differentiable, when $f^{(n)}(t)$ for $0 \leq t < \infty$ exist and $\limsup_{t \to \infty} |f^{(n)}(t)| < \infty$ $(n = 0, 1, \cdots)$. And an Ω_m-function is defined as follows: let $\omega(t)$ be a function defined for $0 \leqq t < \infty$ such that $\omega^{(n)}(t)$ exists for all t and n. Let $\omega^{(j)}(0) = 0$, $0 \leqq j < m$ and $\omega^{(m)}(0) \neq 0$, furthermore $\omega'(t) > 0$ for $t > 0$ and thereby $\omega(t)$ is not bounded in $[0, \infty]$, and thus it is not unlimitedly differentiable. Let. finally, the first derivative of the inverse function of $\sqrt[m]{\omega(t)}$ $(m > 0, t > 0)$ in $[0, \infty]$ be unlimitedly differentiable.

We may write in the following way Wintner's Main Theorem. *Let $\Psi(t)$ be an unlimitedly differentiable function in $[0, \infty]$ and*

[8] Loc. cit., p. 30.

[9] A. Wintner, "Untersuchungen über Funktionen grosser Zahlen," *Math. Zeitschrift*, 28 (1928), pp. 417–429.

let $\omega(t)$ be an Ω_m-function, so that the function $\Phi(t)$ defined by

$$\Phi(t) = \Psi(\omega_m^{-1}(t)) \frac{d}{dt}\omega_m^{-1}(t)$$

is unlimitedly differentiable in $[0, \infty]$; write M_n and $M_{n\epsilon}$ for the upper limits of $|\Phi^{(n)}(t)|$ and $|\Phi_\epsilon^{(n)}(t)|$ in $[0, \infty]$, respectively. Then we get the inequality for $\sqrt[m]{x} \geqq 1$,

$$(3) \quad \left| \int_0^1 e^{-x\omega(t)}\Psi(t)dt - \sum_{\nu=0}^{n-1} \frac{\alpha_{\nu+1}^{(m)}\Phi^{(\nu)}(0)}{(\sqrt[m]{x})^{\nu+1}} \right| \leqq \frac{\Lambda_n}{(\sqrt[m]{x})^{n+1}} \qquad (n = 1, 2, \cdots),$$

where

$$\Lambda_n = \alpha_{n+1}^{(m)} M_n + \frac{(n+1)!}{[\omega(1)]^{n+1}} \left(M_{n1} + \sum_{\nu=0}^{n-1} |\Phi_1^{\nu}(0)| \right),$$

$$(4) \quad \Phi^{(n)}(0) = \left(\frac{d^n}{dt^n} \frac{\Psi(t)t^{n+1}}{\sqrt[m]{(\omega(t))^{n+1}}} \right)_{t=0},$$

and

$$(5) \quad \alpha_n^{(m)} = \frac{\Gamma\left(\frac{n}{m} + 1\right)}{n!}.$$

His inverse problem reads as follows: *Let $\Psi(t)$ and $\omega(t)$ be defined in $[0, 1]$ as unlimitedly differentiable functions such that $\omega^{(j)}(0) = 0$, $0 \leqq j < m$; $\omega^{(m)}(0) \neq 0$; furthermore, $\omega'(t) > 0$ for $0 < t \leqq 1$. Then there is a sequence $\{\Lambda_n\}$ such that the inequality (3) is satisfied for all $n \geqq 1$ and some for $x > 1$, when the coefficients are given by (4) and (5).*

The discussion of the moment integral of the type mentioned at the beginning of this section reduces essentially to that of

$$\int_0^1 e^{-xt}\Phi(t)dt.$$

Let us consider

$$\int_0^\infty t^\lambda e^{-xt}dt = \frac{1}{x^{\lambda+1}} \int_0^\infty y^\lambda e^{-y}dy = \frac{\Gamma(\lambda+1)}{x^{\lambda+1}},$$

which vanishes algebraically at ∞. Then if we write

$$\Phi(t) = \sum a_{\lambda_n} t^{\lambda_n} + R(t),$$

we obtain

$$(6) \qquad \int_0^\infty \Phi(t)e^{-xt}dt = \sum a_{\lambda_n} \frac{\Gamma(\lambda_n + 1)}{x^{\lambda_n+1}} + \int_0^\infty R(t)e^{-xt}dt.$$

Suppose that we determine $\Phi(t)$ such that at the origin

$$g(t) = R(t)t^{-\sigma} \quad \begin{cases} (1) \text{ is bounded.} \\ (2) \text{ tends to zero as } t \to 0. \end{cases}$$

Thus instead of repeated integrations by parts what we do is to subtract a polynomial $\sum a_n t^n$ from Φ and to discuss its remainder.

Now

$$(7) \qquad \int_0^\infty e^{-xt}R(t)dt = \int_0^\infty g(t)e^{-xt}t^\sigma dt$$

$$= \frac{1}{x^{\sigma+1}} \int_0^\infty g\left(\frac{\xi}{x}\right) e^{-\xi}\xi^\sigma d\xi = o(x^{-\sigma-1}),$$

where xt is replaced by ξ.

The kernel of (7), $e^{-\xi}\xi^\sigma$, which we call $K_1(\xi)$, is $O\left(\frac{1}{\xi^2}\right)$ at $\xi = \pm \infty$ and is continuous.

Writing $e^{-\xi} = x$,

$$\int_0^\infty e^{-\xi}e^\sigma e^{iu\xi}d\xi$$

reduces to

$$-\sigma \int_1^0 x^{-iu} \log x dx = \frac{\sigma}{(-iu + 1)^2},$$

which has no zero for $-\infty < u < \infty$. Moreover,

$$\int_0^\infty g\left(\frac{\xi}{x}\right) e^{-\xi}\xi^\sigma d\xi$$

is an absolutely convergent integral. The $K_1(\xi)$ satisfies the conditions required for K_1 in the Key Theorem. We can con-

clude then that

$$\lim_{x\to 0}\int_0^1 g\left(\frac{\xi}{x}\right)K_2(\xi)d\xi = W\int_0^1 K_2(\xi)d\xi = W, \tag{8}$$

in which K_2 of the Key Theorem is 1.

We state the inverse problem as follows: *Let K_2, K_1, and g be functions which satisfy the conditions imposed by the Key Theorem. Let*

$$\lim_{x\to 0}\int_0^1 g\left(\frac{\xi}{x}\right)K_2(\xi)d\xi = W\int_0^1 K_2(\xi)d\xi, \tag{9}$$

$$\int_0^1 g\left(\frac{\xi}{x}\right)K_2(\xi)d\xi \text{ is bounded } [0 < x < \infty].$$

Then

$$\lim_{x\to 0}\int_0^\infty g\left(\frac{\xi}{x}\right)K_1(\xi)d\xi = W\int_0^\infty K_1(\xi)d\xi.$$

If we have

$$K_2(t) = 1, \qquad K_1(t) = e^{-t}t^\sigma, \qquad g(t) = R(t)t^{-\sigma} = o(1),$$

the inverse theorem enables us to write

$$\int_0^\infty \Phi(t)e^{-st}dt = \sum a_{\lambda_n}\frac{\Gamma(\lambda_n+1)}{s^{\lambda_n+1}} + o(s^{-\sigma-1}). \tag{10}$$

Essential points in our method are that we do not assume the existence of the derivatives except at the origin, that we can avoid integrations by parts, and that we need, instead of Taylor's series, only the order of a polynomial near the origin.

ON INTEGRAL FUNCTIONS WITH REAL NEGATIVE ZEROS

By Shikao Ikehara

1. It is well known that the density of the zeros of an integral function is related to the rate of increase of the modulus of the function. In the case of functions of special kinds it takes a very precise form, for instance,

$$f(z) = \prod_{n=1}^{\infty} \left(1 + \frac{z}{a_n}\right),$$

where the numbers $a_1, a_2, \cdots$ are all real and positive, and the product is absolutely convergent; and the asymptotic behaviour assumed is what Valiron has called perfectly regular growth.[1] It may be supposed that $a_{n+1} \geq a_n$ for all values of n. Let $n(r)$ be the number of numbers a_n for which $a_n \leq r$.

The purpose of this paper is to study the problem considered by Titchmarsh [2] from the standpoint of the Tauberian theorems developed by N. Wiener,[3] which are also associated with the names of Bochner,[4] Hardy, Jacob,[5] and Littauer.[6]

[1] G. Valiron, *Integral Functions* (Toulouse, 1923), pp. 44–45.

[2] E. C. Titchmarsh, "On Integral Functions with Real Negative Zeros," *Proc. London Math. Soc.*, Ser. 2, 26 (1926), pp. 185–200.

[3] N. Wiener, "On a Theorem of Bochner and Hardy," *Jour. London Math. Soc.*, 2 (1927), pp. 118–123; "A New Method in Tauberian Theorem," Jour. of Math. and Phys., Massachusetts Inst. Tech., 7 (1928), pp. 161–184; "A Type of Tauberian Theorem Applying to Fourier Series," *Proc. London Math. Soc.*, Ser. 2, 30 (1929), pp. 1–8; "Generalized Harmonic Analysis," *Acta Mathematica*, vol. 55 (1930), pp. 117–258.

[4] Bochner and Hardy, "Notes on Two Theorems of Norbert Wiener," *Jour. London Math. Soc.*, 1 (1926), pp. 240–244.

[5] M. Jacob, "Über ein Theorem von Bochner-Hardy-Wiener," *Jour. London Math. Soc.*, 3 (1928), pp. 182–187.

[6] S. B. Littauer, "Note on a Theorem of Jacob," *Jour. London Math. Soc.*, 4 (1929), pp. 226–231. "A New Tauberian Theorem with Application to the Summability of Fourier Series and Integrals," Jour. of Math. and Phys., 8 (1929), pp. 216–234.

Theorems of Titchmarsh are as follows:

THEOREM I_T. *If, as* $r \to \infty$,

$$n(r) \sim \lambda r^{\rho} \quad (0 < \rho < 1, \lambda > 0), \tag{1.1}$$

then, as $x \to \infty$ (*x real*),

$$\log f(x) \sim \pi\lambda \operatorname{cosec} \pi\rho \cdot x^{\rho}. \tag{1.2}$$

THEOREM II_T. *If, as* $x \to \infty$,

$$\log f(x) \sim \pi\lambda \operatorname{cosec} \pi\rho \cdot x^{\rho} \cdot (0 < \rho < 1, \lambda > 0), \tag{1.3}$$

then

$$n(r) \sim \lambda r^{\rho}. \tag{1.4}$$

THEOREM III_T. *Let*

$$n(x) \sim \lambda x^{\rho} (0 < \rho < 1). \tag{1.5}$$

Then, if ϵ *is any given positive number,*

$$\log |f(-x)| > (\pi\lambda \cot \pi\rho + \epsilon)x^{\rho} \quad [x > x_0(\epsilon)],$$

while, if η *is any other positive number,*

$$\log |f(-x)| > (\pi\lambda \cot \pi\rho - \epsilon)x^{\rho},$$

for $0 < x < X$, $\{X > X_0(\epsilon, \eta)\}$, *except in a set of measure less than* ηX.

We may express the result by saying that

$$\log |f(-x)| \sim \pi\lambda \cot \pi\rho \cdot x^{\rho} \tag{1.6}$$

in a set of density 1.

THEOREM IV_T. *If*

$$\log |f(-x)| < (\pi\lambda \cot \pi\rho + \epsilon)x^{\rho} (x > x_0) \tag{1.71}$$

and

$$\log |f(-x)| > (\pi\lambda \cos \pi\rho - \epsilon)x^{\rho} \tag{1.72}$$

in a set of density 1, and if also $n(x) = o(\sqrt{x})$, *then*

$$n(x) \sim \lambda x^{\rho}. \tag{1.8}$$

2. Closely related to these theorems we shall prove the following theorems:

THEOREM I. *If*

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T\frac{n(r)}{r^\rho}dr=A,\quad(0<\rho<1,\lambda>0) \tag{2.1}$$

then

$$\lim_{x\to\infty}\int_0^\infty\frac{x^{1-\rho}n(r)}{r(x+r)}dr=A\int_0^\infty\frac{r^{\rho-1}}{1+r}dr. \tag{2.2}$$

The conclusion (2.2) is equivalent to (1.2) in Theorem I_T if we notice that:[7]

$$\begin{aligned}\log f(x)&=\sum_{n=1}^{\infty}\log\left(1+\frac{x}{a_n}\right)\\&=\sum_{n=1}^{\infty}n\left\{\log\left(1+\frac{x}{a_n}\right)-\log\left(1+\frac{x}{a_{n+1}}\right)\right\}\\&=\sum_{n=1}^{\infty}n\int_{a_n}^{a_{n+1}}\frac{xdr}{r(x+r)}\\&=x\int_0^\infty\frac{n(r)dr}{r(x+r)}.\end{aligned}$$

Hence

$$\begin{aligned}\log f(x)\sim x\int_0^\infty\frac{\lambda r^\rho dr}{r(x+r)}&=\lambda x^\rho\int_0^\infty\frac{t^{\rho-1}}{1+t}dt\\&=\pi\lambda\operatorname{cosec}\pi\rho\cdot x^\rho.\end{aligned}$$

The converse of this theorem is also true.

THEOREM II. *If,*

$$\lim_{x\to\infty}\int_0^\infty\frac{x^{1-\rho}n(r)dr}{r(x+r)}=A\int_0^\infty\frac{r^{\rho-1}}{1+r}dr, \tag{2.3}$$

then

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T\frac{n(r)}{r^\rho}dr=A. \tag{2.4}$$

Functions of the type considered in the two theorems above are characterized by their asymptotic behaviour on the positive

[7] E. C. Titchmarsh, ibid., pp. 185–186.

real axis. They are also, to a certain extent, characterized by their asymptotic behaviour on the negative real axis.

THEOREM III. *Let*

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T\frac{n(r)dr}{r^\rho}=A\quad(0<\rho<1),\tag{2.5}$$

then,

$$\lim_{x\to\infty}\int_0^\infty\left|\frac{-x^{1-\rho}n(r)}{r(-x+r)}\right|dr=A\int_0^\infty\frac{r^{\rho-1}}{1+r}dr\tag{2.6}$$

in a set of density 1.

THEOREM IV. *If,*

$$\lim_{x\to\infty}\int_0^\infty\left|\frac{-x^{1-\rho}n(r)}{r(-x+r)}\right|dr=A\int_0^\infty\frac{r^{\rho-1}}{1+r}dr\tag{2.7}$$

in a set of density 1, *then*

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T\frac{n(r)}{r^\rho}dr=A.\tag{2.8}$$

In order to prove our theorems we depend solely on the Tauberian theorem of Wiener,[8] which, in one form, reads:

KEY THEOREM: *Let* $M_1(\xi)$ *and* $M_2(\xi)$ *be two bounded functions which are* $O(1/\xi^2)$ *at* $\neq\infty$. *Let* $M_1(\xi)$ *be measurable and non-negative, and let*

$$\int_{-\infty}^{\infty}M_1(\xi)e^{iu\xi}d\xi\not\equiv 0\quad[-\infty<u<\infty].$$

Let $M_2(\xi)$ *be continuous, except for a finite number of finite jumps. Let* $g(\xi)$ *be a non-negative measurable function such that*

(*a*) $$\lim_{\eta\to\infty}\int_{-\infty}^{\infty}g(\eta-\xi)M_1(\xi)d\xi=C\int_{-\infty}^{\infty}M_1(\xi)d\xi;$$

(*b*) $$\int_{-\infty}^{\infty}g(\eta-\xi)M_1(\xi)d\xi \text{ is bounded}\quad[-\infty<\eta<\infty].$$

[8] N. Wiener, "A New Method in Tauberian Theorem," JOUR. OF MATH. AND PHYS., MASSACHUSETTS INST. TECH., 7 (1928), pp. 161–184.

Then

$$(c) \qquad \lim_{\eta \to \infty} \int_{-\infty}^{\infty} g(\eta - \xi) M_2(\xi) d\xi = C \int_{-\infty}^{\infty} M_2(\xi) d\xi.$$

In the Theorem I we write $r = x/y$; $T = 1/y$; $x = e^{\xi}$; $y = e^{\eta}$, then (2.1) can be written as

$$(2.11) \qquad \lim_{\eta \to -\infty} \int_{-\infty}^{0} \left[\frac{n(e^{\xi-\eta})}{e^{\rho(\xi-\eta)}} \right] e^{\xi} d\xi = A \int_{-\infty}^{0} e^{\xi} d\xi.$$

where $\left[\frac{n(e^{\xi-\eta})}{e^{\rho(\xi-\eta)}} \right]$ and e^{ξ} correspond, respectively, to $g(\eta - \xi)$ and $M_1(\xi)$ in the Key Theorem. And (2.2) is by similar transformations,

$$(2.21) \qquad \lim_{\eta \to -\infty} \int_{-\infty}^{\infty} \left[\frac{n(e^{\xi-\eta})}{e^{\rho(\xi-\eta)}} \right] \left(\frac{e^{\rho\xi}}{1 + e^{\xi}} \right) d\xi = A \int_{-\infty}^{\infty} \frac{e^{\rho\xi}}{1 + e^{\xi}} d\xi.$$

The proof follows almost directly from the Key Theorem upon showing that the conditions of that theorem are fulfilled.

The kernel of (2.11) is e^{ξ} for $-\infty \leq \xi \leq 0$, and zero for $0 < \xi \leq \infty$. At $\xi = \neq \infty$, it is $O(1/\xi^2)$; furthermore, e^{ξ} is continuous. Since

$$\log(x) \geq x \int_{x}^{\infty} \frac{n(x)dr}{r(x + r)} > xn(x) \int_{x}^{\infty} \frac{dr}{r(x + r)} = n(x) \log 2,$$

we have [9]

$$n(x) = o(x^{\rho}),$$

or, in our notation,

$$\frac{n(e^{\xi-\eta})}{e^{\rho(\xi-\eta)}} = O(1),$$

and it is a non-negative measurable function.

Since, then, $\int_{-\infty}^{\infty} \frac{n(e^{\xi-\eta})}{e^{\rho(\xi-\eta)}} d\xi$ is bounded, and since e^{ξ} is bounded and $O(1/\xi^2)$ at $\neq \infty$, independent of η,

$$(2.12) \qquad \int_{-\infty}^{0} \left[\frac{n(e^{\xi-\eta})}{e^{\rho(\xi-\eta)}} \right] e^{\xi} d\xi$$

is bounded for $-\infty < \eta < \infty$.

[9] E. C. Titchmarsh, ibid., p. 187.

Moreover,

$$\int_{-\infty}^{0} e^{\xi} e^{iu\xi} d\xi$$

has no zeros for $-\infty < u < \infty$. The kernel of (2.21) is obviously

$$\frac{e^{\rho\xi}}{1+e^{\xi}} = O(1/\xi^2) \quad (0 < \rho < 1)$$

Let us now form, as in Wiener's [10] paper,

$$K_2(\xi) = \int_{-\infty}^{\infty} K_1(\eta) R(\xi - \eta) d\eta \tag{2.3}$$

in which $K_1(\xi) = e^{\xi}$ in (2.11) and $\int_{-\infty}^{\infty} |R(\eta)| d\eta$ converges and for which (2.21) is a consequence of (2.11) and (2.12) for a non-negative measurable function,

$$f(\xi) = \frac{n(e^{\xi})}{e^{\rho\xi}}.$$

Since the double integral

$$\int_{-\infty}^{\infty} R(\zeta) d\zeta \int_{-\infty}^{\infty} f(\xi - \eta + \zeta) K_1(\xi) d\xi$$

is absolutely convergent,

$$\begin{aligned} \int_{-\infty}^{\infty} K_2(\xi) f(\xi - \eta) d\xi &= \int_{-\infty}^{\infty} f(\xi - \eta) d\xi \int_{-\infty}^{\infty} K_1(\xi - \zeta) R(\zeta) d\zeta \\ &= \int_{-\infty}^{\infty} R(\zeta) d\zeta \int_{-\infty}^{\infty} f(\xi - \eta + \zeta) K_1(\xi) d\xi, \end{aligned}$$

and

$$\begin{aligned} \lim_{\eta \to -\infty} \int_{-\infty}^{\infty} K_2(\xi) f(\xi - \eta) d\xi &= \lim_{\eta \to -\infty} \int_{-\infty}^{\infty} R(\zeta) d\zeta \int_{-\infty}^{\infty} f(\xi - \eta + \zeta) K_1(\xi) d\xi \\ &= A \int_{-\infty}^{\infty} R(\zeta) d\zeta \int_{-\infty}^{\infty} K_1(\xi) d\xi \\ &= A \int_{-\infty}^{\infty} R(\zeta) d\zeta \int_{-\infty}^{0} e^{\xi} d\xi = A \int_{-\infty}^{\infty} K_2(\xi) d\xi. \end{aligned}$$

[10] N. Wiener, reference 8.

Hence

$$R(\xi) = K_2(\xi) = \frac{e^{\rho\xi}}{1 + e^{\xi}}$$

which is absolutely integrable. In general cases, (2.3) can be solved by means of the Plancherel [11] theory of the Fourier transform as shown by Wiener. This completes the proof.

In Theorem II we merely point out that:

$$\int_{-\infty}^{\infty} \frac{e^{\rho\xi} e^{iu\xi}}{1 + e^{\xi}} d\xi \neq 0 \quad (-\infty < u < \infty ; 0 < \rho < 1)$$

and that, from the previous consideration of $f(\xi)$ and $K_1(\xi)$,

$$\int_{-\infty}^{\infty} \left[\frac{n(e^{\xi-\eta})}{e^{\rho(\xi-\eta)}}\right]\left[\frac{e^{\rho\xi}}{1 + e^{\xi}}\right] d\xi$$

is bounded for $-\infty < \eta < \infty$. Thus, Theorem II satisfies the conditions required by the Key Theorem.

By the same method of proof it can be easily shown that Theorems III and IV satisfy the Wiener theorem.

3. It is to be pointed out that Titchmarsh proves Theorem II$_T$ by means of a "positive" Tauberian theorem due to Hardy and Littlewood [12] and, in turn, he deduces the latter theorem from his theorem above.

Theorem of Hardy and Littlewood:

If $F(x)$ is a positive integrable function such that, when $t \to 0$,

$$\int_0^{\infty} F(x)e^{-xt}dx \sim t^{-\beta} \quad (\beta > 0), \tag{3.1}$$

then, when $x \to \infty$,

$$\int_0^{x} F(u)du \sim \frac{x^{\beta}}{\Gamma(\beta + 1)} . \tag{3.2}$$

[11] M. Plancherel, "Contribution à l'étude de la représentation d'une fonction arbitraire par des intégrales définies," *Rend. di Palermo*, 30 (1910), pp. 289–335.

[12] Hardy and Littlewood, "Tauberian Theorems concerning Power Series and Dirichlet's Series whose Coefficients are Positive," *Proc. London Math. Soc.*, Ser. 2, 13 (1913), pp. 174–191.

We can include this theorem in the Tauberian theorem of Wiener. Write $xt = y$; $y = e^{\xi}$; $t = e^{\eta}$ in (3.1), then

$$\lim_{\eta \to -\infty} \int_{-\infty}^{+\infty} \frac{F(e^{\xi-\eta})}{e^{(\beta-1)(\xi-\eta)}} (e^{\beta\xi}e^{-\xi})d\xi = C \int_{-\infty}^{+\infty} (e^{\beta\xi}e^{-\xi})d\xi, \tag{3.11}$$

and (3.2) becomes by substitutions: $u = xy$; $x = e^{-\eta}$; $y = e^{\xi}$,

$$\lim_{\eta \to -\infty} \int_{-\infty}^{+\infty} \frac{F(e^{\xi-\eta})}{e^{(\beta-1)(\xi-\eta)}} e^{\beta\xi}d = C \int_{-\infty}^{+\infty} e^{\beta\xi}d\xi. \tag{3.21}$$

A NOTE ON INFINITE PRODUCTS

By Shikao Ikehara

Let

$$f(x) = \prod_{n=1}^{\infty} \left(1 + \frac{x^2}{a_n^2}\right), \tag{1}$$

where the numbers $a_1, a_2, \cdots$ are all real and positive, and the product is absolutely convergent. It may be supposed that $a_{n+1} \geq a_n$ for all values of n. Let $n(r)$ be the number of numbers a_n for which $a_n \leq r$. Titchmarsh [1] proves:

If $a_n \sim \pi n/\lambda$, *then* $\log f(x) \sim \lambda x$. *And the converse is also true.*

Closely related to the theorem above we have:

If, when $T \to \infty$,

$$\frac{1}{T}\int_0^T \frac{n(r)}{r}\,dr \sim \frac{\lambda}{\pi}, \tag{2}$$

then, when $x \to \infty$,

$$x^{-1} \log f(x) \sim \lambda. \tag{3}$$

And the converse is also true.

By elementary transformations we can write (2) as

$$\lim_{v\to\infty} \int_0^v \left[\frac{n(xy)}{xy}\right] dx = A \int_0^v dx, \tag{4}$$

and (3) as

$$\lim_{v\to\infty} \int_0^\infty \left[\frac{n(xy)}{xy}\right]\left[\frac{1}{1+x^2}\right] dx = A \int_0^\infty \frac{1}{1+x^2}\,dx. \tag{5}$$

The kernel in (4) is 1, and

$$\int_0^\infty 1 \cdot e^{iu} dx \neq 0 \quad [-\infty < u < \infty].$$

[1] E. C. Titchmarsh, "A Theorem on Infinite Products," *Journal London Math. Soc.*, 1 (1926), 1–3.

Since

$$\log f(x) \geqq 2x^2 \int_x^\infty \frac{n(r)dr}{r(x^2+r^2)} > 2x^2 n(x) \int_x^\infty \frac{dr}{r(x^2+r^2)} = n(x) \log 2,$$

$$n(x) = o(x), \tag{6}$$

and hence (4) is bounded for $0 < y < \infty$.

The kernel in (5) is $1/(1+x^2)$, which is $O(1/x(\log x)^2)$ both at 0 and at ∞, and it is also

$$\int_0^\infty \frac{x^{iu}dx}{1+x^2} \neq 0 \quad [-\infty < u < \infty].$$

From the boundedness of the kernel of (5) and (6) it is easy to see the expression of (5) is bounded for $0 < y < \infty$. If we write

$$\begin{cases} M_1(x) = \begin{cases} 1; [0 \leqq x \leqq y], \\ 0; [x > y] \end{cases} \\ M_2(x) = \dfrac{1}{1+x^2}; \\ f(x) = n(x)/x, \end{cases}$$

then our theorem satisfies the conditions required by the Wiener Theorem:[2]

Let $M_1(x)$ and $M_2(x)$ be two functions bounded over every range $(\epsilon, 1/\epsilon)$, which are $O\left(\dfrac{1}{x(\log x)^2}\right)$ at 0 and ∞. Let $M_1(x)$ be measurable and non-negative, and let

$$\int_0^\infty M_1(x)x^{iu}dx \neq 0 \quad [-\infty < u < \infty].$$

Let $M_2(x)$ be continuous, except for a finite number of finite jumps. Let $f(x)$ be a measurable function bounded below and such that

[2] N. Wiener, "A New Method in Tauberian Theorem," JOURNAL MATH. AND PHYS., MASS. INST. OF TECH., 7 (1928), 161–184.

$$\lim_{y \to 0[\infty]} \int_0^\infty f(xy) M_1(x) dx = C \int_0^\infty M_1(x) dx;$$

$$\int_0^\infty f(xy) M_1(x) dx \text{ is bounded } [0 < y < \infty].$$

Then

$$\lim_{y \to 0[\infty]} \int_0^\infty f(xy) M_2(x) dx = C \int_0^\infty M_2(x) dx.$$

Here ∞ is put into brackets to indicate that at these points it may be consistently substituted for 0.

PFAFFIANS IN PARAMETRIC FORM[1]

By C. L. E. Moore and P. Franklin

1. Introduction. In certain previous work on the geometric interpretation of Pfaffian equations,[2] the parametric representation of surfaces has been modified to give a parametric representation of Pfaffians. No precise discussion was made as to what classes of Pfaffians could be parametrically represented, and no restrictions on the defining equations which had to hold to yield a Pfaffian were mentioned. In the present paper we investigate the possibility of representing the differentials of the coördinates for a Pfaffian in three space in terms of two parameters, and find that if such a representation is possible when Cartesian coördinates are used, the elements of the Pfaffian are unchanged by a certain translation, at least in the non-integrable case. When such a representation is possible for a non-integrable Pfaffian as well as for its arc element when curvilinear coördinates are used, the elements of the Pfaffian are unchanged by a certain helical motion of space. Furthermore, the defining relations may not be taken at random, but must satisfy some stringent conditions, which are obtained. Thus the method of parametric representation, in its application to Pfaffians, is relatively restricted in scope.

All Pfaffians with a parametric representation have elements of arc which can be expressed in terms of two parameters. We show, however, that they are not the most general class of Pfaffians with a two-parameter arc element, and give a geometric characterization of the latter class.

2. Pfaffians Parametric with Cartesian Coördinates. A

[1] Presented to the American Mathematical Society.

[2] Abbé Issaly, "Pseudo-Surface," *Bulletin de la Société Math. de France*, 1888; *Nouvelles Annales*, 1900 and 1901.

pseudo-surface has been defined by a set of equations of the form

$$
\begin{aligned}
dx &= X_1 du + X_2 dv, \\
dy &= Y_1 du + Y_2 dv, \\
dz &= Z_1 du + Z_2 dv,
\end{aligned} \tag{1}
$$

in which the X_i, Y_i, Z_i are functions of the two parameters u and v. It is assumed that there is a pair of parameters u, v uniquely determined for each point x, y, z. The nature of this correspondence has not been precisely explained. Presumably it is continuous, and arises from relations supplementing (1).

A simple example of a pseudo-surface equivalent to a Pfaffian is

$$
\begin{aligned}
dx &= du, \\
dy &= dv, \\
dz &= v du - u dv.
\end{aligned} \tag{2}
$$

If, for any one point, x_0, y_0, z_0 we have the supplementary relation

$$u_0 = x_0 + a, \qquad v_0 = y_0 + b, \tag{3}$$

we may deduce from (2) and continuity considerations that

$$u = x + a, \qquad v = y + b \tag{4}$$

and hence conclude that (2) is equivalent to the Pfaffian

$$dz = (y + b)dx - (x + a)dy, \tag{5}$$

which corresponds to a null system.

Let us next consider a general set of equations (1), together with some supplementary relations which when combined with (1) determine a continuous correspondence of pairs (u, v) and triplets (x, y, z) necessarily for the most part one to infinity. Such a system will determine a Pfaffian, since at each point x, y, z the differentials satisfying (1) for the corresponding fixed u, v, but for arbitrary du and dv give rise to a planar element.

The single infinity of points corresponding to a given pair of parameters (u, v) will in general constitute a curve. We define a new parameter s, as the arc length measured along these curves, starting from their intersection with some arbitrary fixed surface

cutting all of them. We thus have a correspondence between (x, y, z) and (u, v, s) which is continuous and in general one to one. Now consider a simple closed curve in the u, v plane, beginning and ending at u_0, v_0. If we select s_0 arbitrarily, the point (u_0, v_0, s_0) will yield a point (x_0, y_0, z_0). As we move around the closed u, v curve, the new values of x, y, z and hence s will be determined by continuity. There are now two possibilities. Either (a) the curve in x, y, z space will be closed for all choices of the u, v curve, or (b) for some u, v curve the final value of x, y, z will be distinct from x_0, y_0, z_0. We consider these two cases separately.

(a) Since the final values of x, y, z are the same as x_0, y_0, z_0, the right members of (1) give zero when integrated over any closed circuit, and hence are exact differentials. They may accordingly be integrated, giving in place of (1):

$$x = a + f(u, v),$$
$$y = b + g(u, v),$$
$$z = c + h(u, v). \tag{6}$$

If a, b, c are unrestricted, this system of equations represents the triple infinity of surfaces obtained from a single one by all possible translations. The supplementary relations which make the determination of u, v by x, y, z unique must give two equations

$$F(a, b, c) = 0,$$
$$G(a, b, c) = 0, \tag{7}$$

restricting these translations to a single infinity. Thus, in this first case the system of equations (1) together with the supplementary relations are equivalent to an integrable Pfaffian. The integral surfaces of this Pfaffian only differ from one another by rigid displacements, corresponding to chords of the curve (7).

(b) Since the final point x, y, z is distinct from x_0, y_0, z_0 the final value of s will differ from s_0. This final value will depend on s_0 and the closed curve. It will, in general, vary if we keep s_0 fixed and vary the closed curve. In particular, if we shrink

the curve to a point in a continuous manner, the final value of s will change continuously and approach s_0. Thus this final value may be made any one between that first reached and s_0. But, we may move around any of the u, v circuits several times, and may reverse the sense of description. Thus we may by a suitable choice of the circuit, make the final value of s any whatever. Next consider the curve (γ) of points in x, y, z space corresponding to a particular u_0, v_0, but all possible values of s_0 and their displacements in space for a particular circuit through u_0, v_0. Since du and dv are prescribed by the given circuit, by (1) the differential vector dx, dy, dz will be the same for all points of the curve (γ), so that the motion of this curve (γ) will be a translation. Thus this curve (γ) must be such that it is possible to bring the curve into itself by a translation, which may take any point on it into any other point. That is the curve (γ) of points for any fixed u_0, v_0 must be a straight line. Since a curve joining two distinct values of u, v in the u, v plane determines in space a parallel displacement of the corresponding straight lines in x, y, z space, these straight lines are all parallel. We may accordingly change the coördinates in such a way that these lines are all parallel to the z axis. Then the first two expressions in (1) will be exact, and a change of parameters will make $x = u$, $y = v$. The system of equations will now be:

$$
\begin{aligned}
(8) \qquad dx &= du, \\
dy &= dv, \\
dv &= M(u, v)du + N(u, v)dv.
\end{aligned}
$$

In view of the supplementary relations

$$(9) \qquad x_0 = u_0, \qquad y_0 = v_0.$$

these are equivalent to the Pfaffian

$$(10) \qquad dz = M(x, y)dx + N(x, y)dy.$$

Thus the example (2) above is typical of this case.[3]

[3] For the rôle of this type of Pfaffian in the canonical form of Pfaffians analogous to quadrics, cf. this JOURNAL, vol. X, no. 1, p. 32, eq. (5.31) through (5.36).

We summarize our results in

THEOREM I. *If a Pfaffian equation with coefficients functions of x, y, z:*

$$Adx + Bdy + Cdz = 0,$$

is equivalent to a system of the form

$$\begin{aligned} dx &= X_1du + X_2dv, \\ dy &= Y_1du + Y_2dv, \\ dz &= Z_1du + Z_2dv. \end{aligned}$$

with coefficients functions of the parameters u, v, together with supplementary relations which determine a unique pair u, v for each point x, y, z, the correspondence being continuous, then either the Pfaffian is integrable, and the integral

$$f(x, y, z) = k$$

represents a family of surfaces which can all be obtained from one by the parallel displacements which follow a certain space curve, or, by a suitable change of coördinates, the Pfaffian may be reduced to:

$$dz = M(x, y)dx + N(x, y)dy.$$

It should be noted that this theorem places heavy restrictions on the functions X_i, Y_i, Z_i of (1), when such a system arises from a Pfaffian. For, consider a linear combination with constant coefficients

$$a_1dx + a_2dy + a_3dz, \tag{11}$$

expressed in terms of u and v by means of (1). In case (a) every such combination is exact, since the separate terms are. In case (b) there must be two linearly independent such combinations to serve as differentials of the transformed x and y. Since the condition for exactness is the vanishing of the linear combination of

$$\partial X_1/\partial v - \partial X_2/\partial u, \qquad \partial Y_1/\partial v - \partial Y_2/\partial u, \qquad \partial Z_1/\partial v - \partial Z_2/\partial u$$

analogous to (11), these quantities must be proportional. That is, there must exist three constants, not more than one zero, such that

$$k_1(\partial X_1/\partial v - \partial X_2/\partial u) = k_2(\partial Y_1/\partial v - \partial Y_2/\partial u) = k_3(\partial Z_1/\partial v - \partial Z_2/\partial u).$$

Conversely, when this condition holds, we may find two linearly independent combinations (11) which are exact, and hence two corresponding to perpendicular directions. On taking these directions as the new x and y directions and making the corresponding change in parameters we shall have the canonical form for b) unless all *three* of the original expressions are *exact*. In this event, we have case (a) and the constants of integration in the first two expressions *need not* be taken the same for all points.

3. Pfaffians Parametric with Curvilinear Coördinates, Having Parametric Arc Elements. In the type of Pfaffian just discussed, the arc element, as well as the differentials of the coördinates, can be expressed in terms of the two parameters. This results from the fact that the Euclidean arc element only involves the differentials of the Cartesian coördinates. Let us now consider a Pfaffian such that the arc element can be expressed in terms of two parameters, and also the differentials of a properly chosen system of curvilinear coördinates can be expressed in terms of the same two parameters.

Denote by u, v the parameters in terms of which the arc element is expressed. A given pair of values u, v will yield a curve in space. If we take w any parameter along this curve, we will have a correspondence between x, y, z and u, v, w in general one to one, so that we may take u, v, w as curvilinear coördinates. Since the arc element is determined by u, v and their differentials, but w is not determined by u, v; the arc element in terms of the curvilinear coördinates can not involve w explicitly. For arbitrary elements in space, the arc element involves dw. For elements in the Pfaffian, dw is given in terms of u, v and their differentials, and so may be eliminated. The system of equations

analogous to (1), or rather (8) for this case is:

$$(12) \qquad \begin{aligned} du &= du, \\ dv &= dv, \\ dw &= M(u, v)du + N(u, v)dv. \end{aligned}$$

From the fact that the arc element is independent of w, the transformation of space:

$$(13) \qquad u' = u, \qquad v' = v, \qquad w' = w + a$$

gives a rigid displacement of space, since distances are preserved. Hence the path curves of this displacement are helices. In other words, the curves u = constant, v = constant which are taken into themselves by this displacement are helices. Since these helices are all generated by the same infinitesimal displacement, they have a common axis, which we may take as the axis of z, and a common pitch which we denote by tan k. After making a suitable change in the parameters, if necessary, we may then write:

$$(14) \qquad \begin{aligned} du &= dr, \\ dv &= \cos kdz - \sin kd\theta, \\ dw &= \sin kdz + \cos kd\theta. \end{aligned}$$

These equations express the curvilinear coördinates u, v, w in terms of the polar coördinates z, r, θ. We have for the arc element

$$(15) \qquad \begin{aligned} ds^2 &= dr^2 + dz^2 + r^2d\theta^2 \\ &= du^2 + (\cos kdv + \sin kdw)^2 + (-\sin kdv + \cos kdw)^2. \end{aligned}$$

Any Pfaffian which becomes

$$(16) \qquad dw = M(u, v)du + N(u, v)dv$$

when expressed in terms of the coördinates given by (14) will lead to an arc element expressible in terms of u and v by (15). The discussion shows that all Pfaffians whose coördinates, and arc elements can have their differentials expressed in terms of two parameters are of this type. We express this in

THEOREM II. *If the arc element of a Pfaffian is expressible in terms of two parameters which may be taken with some third coördinate to form a system of curvilinear coördinates for space, whose differentials are expressible in terms of the parameters, the elements of the Pfaffian may be taken into themselves by a family of helical displacements.*

The simplest example of this situation is that of Theorem I, where the helices reduced to straight lines. Another simple example is that in which they reduce to circles with a common axis. The Pfaffian (2) may be written

$$dz = -r^2 d\theta \tag{17}$$

in terms of polar coördinates, and so is an example of this case.[4]

4. Pfaffians with Arc Elements Expressible Parametrically. There are Pfaffians such that all elements of arc in them are expressible in terms of two parameters distinct from the type just discussed. For these the parameters in question may not be linked up with curvilinear coördinates.

We begin with the integrable case. If u, v are the parameters, and the Pfaffian is integrable, these two quantities may be used as the parameters for any one of the surfaces forming the Pfaffian. This follows from the fact that if we start with any point on a particular integral surface, and vary u, v in all possible manners, we will always remain in the surface as long as we remain in the Pfaffian. Conversely, any real curve in the surface (hence not a minimal line) will have to be such that no two distinct points on it have the same values of u and v in order to make the arc length different from zero. Since the arc element has the same form in u and v for all the surfaces, they are all applicable to one another, and we have:

THEOREM III. *When the arc element of an integrable Pfaffian can be expressed in terms of two parameters, the integral surfaces of the Pfaffian are all applicable to one another.*

[4] Similar examples are the canonical forms (5.21), (5.22), (5.26), (5.31), (5.34), this JOURNAL, vol. X, no. 1, p. 32.

There are other restrictions on the surfaces, since it must be possible to write their equation in the form

$$f(x, y, z) = c, \tag{18}$$

the value of the constant c determining which surface is in question. This applicability requirement is certainly met in our earlier cases, where the surfaces could be applied to one another by a rigid displacement. To show that there are Pfaffians of the present kind not included in the previous discussion, we give the example:

$$y^3dx - (x^2z + xy^2)dy + x^2ydz = 0. \tag{19}$$

This is an integrable Praffian, the integral being

$$z/y - y/x = c, \tag{20}$$

so that the integral surfaces are a family of cones. If we take u, v as Cartesian coördinates in the (developed) conical surfaces, the arc element for the Pfaffian will be

$$ds^2 = du^2 + dv^2. \tag{21}$$

Except for the requirement of continuity, u and v are to a great extent unrestricted. If we attempt to include the whole of space, we meet difficulties owing to the multiple valued character of the parameters, but in a small portion of space there is no complication. The cones given by (20) are clearly not superimposable by a rigid displacement.

To understand the structure of non-integrable Pfaffians whose arc elements are expressible in terms of two parameters, we observe that for any system of curvilinear coördinates, u_1, u_2, u_3, the Euclidean arc element may be written

$$ds^2 = \sum g_{ij}du_idu_j. \tag{22}$$

Now suppose that, for a particular choice of curvilinear coördinates, three functions of u_1 and u_2 only, a, b, c can be found such

that:

$$\begin{vmatrix} g_{11} - a & g_{12} - b & g_{13} \\ g_{21} - b & g_{22} - c & g_{23} \\ g_{31} & g_{32} & g_{33} \end{vmatrix} = 0. \tag{23}$$

Then, when the form

$$adu_1^2 + 2bdu_1du_2 + cdu_2^2 \tag{24}$$

is subtracted from the expression (22), the remainder is decomposable into two linear factors, so that we may write:

$$\begin{aligned} ds^2 = {} & adu_1^2 + 2bdu_1du_2 + cdu_2^2 \\ & + (Adu_1 + Bdu_2 + Cdu_3)(Ddu_1 + Edu_2 + Fdu_3). \end{aligned} \tag{25}$$

The capital letters will in general involve all three variables. If we now put either

$$Adu_1 + Bdu_2 + Cdu_3 = 0, \tag{26}$$

or

$$Ddu_1 + Edu_2 + Fdu_3 = 0, \tag{27}$$

we shall have a Pfaffian whose arc element will reduce to (24) and thus be expressible in terms of two parameters.

As one example of such a situation, we consider the Pfaffian (19). We take u_1 and u_2 the u and v of (21), and u_3 any third coördinate such that du_3 is not normal to the surfaces (20). Then, since $du_3 = 0$ reduces the arc element to (21), the element for these curvilinear coördinates may be written:

$$ds^2 = {}'du_1^2 + du_2^2 + (Adu_1 + Bdu_2 + Cdu_3)du_3. \tag{28}$$

Presumably for some choice of the parameters u_1, u_2 and u_3 the Pfaffian (26) with coefficients taken from (28) will be non-integrable.

The pairing up of Pfaffians of this type may be anticipated from geometric considerations. For, geometrically, we may start with *any* congruence of curves, and select a point on one of the curves as well as a planar element through this point. A small

circle in this plane will determine a "tube" formed of nearby curves. At any other point on the curve we may draw a sphere of radius equal to that of the circle determining the tube, and its intersection with the tube will be two curves, which give rise to *conical* surfaces with vertex at the center of the sphere. In the case of interest to us, one of these conical surfaces must reduce to a plane for all points in space. But, in this case, the tube is in the neighborhood of any point, to the first order of infinitesimals, a right elliptic cylinder. Consequently both the sections made by the sphere will be plane curves and their planes will make equal angles with the curves of the congruence. It is intuitively clear that one set of elements may be those of an integrable surface, while the other is not. This was the case in the example deduced from (28). The congruence we start with will have to be very special in order that all the cones reduce to planes, and we have a Pfaffian of the kind sought.

THE MAPPING THEOREM OF KOEBE AND THE PROBLEM OF PLATEAU

By Jesse Douglas

In a recent paper [1] the writer gave the first solution of the problem of Plateau for an arbitrary Jordan contour, proving that every Jordan curve in n-dimensional euclidean space is the boundary of a simply-connected portion of a minimal surface. In this treatment the conformal mapping of a plane Jordan region on the disc of the unit circle was included as the special case $n = 2$, and thereby a new and in many respects clearer and more elementary proof was given for the Riemann mapping theorem (concerning the conformal mapping of the interiors) together with the supplementary theorem of Osgood and Carathéodory (asserting the one-one continuous mapping of the boundaries). Since it was an essential part of the purpose of the paper to synthesize the theory of conformal mapping with that of minimal surfaces, it was indispensable not to assume anything of the former theory but to proceed directly from the axioms of analysis. This was done, and a simultaneous treatment of the two theories developed which was elementary and completely self-sufficient.

After the existence of the minimal surface had been established, we used a theorem on the conformal mapping of polyhedra from the theory of conformal mapping as developed by Koebe [2] in a brief proof of the least area property. The present paper uses

[1] "Solution of the Problem of Plateau," *Transactions of the American Mathematical Society*, vol. 33 (1931), no. 1, pp. 262–321. The phrase "the cited paper" whenever used hereafter will refer to this one.

[2] P. Koebe, "Allgemeine Theorie der Riemannschen Mannigfaltigkeiten," *Acta Mathematica*, vol. 50 (1927), pp. 119 et seq. This theorem was first used in connection with the problem of Plateau by T. Radó, *Annals of Mathematics*, vol. 31 (1930), pp. 457–469, and *Mathematische Zeitschrift*, vol. 32 (1930), pp. 762–796.

this conformal mapping of a polyhedral surface to give an alternative proof of the existence of the minimal surface. The proof amounts to adjoining to the polyhedral mapping theorem certain supplementary considerations used in another way in the cited paper. While it requires a certain amount of time and space to set down these considerations explicitly, they are of so simple and evident a nature as to justify considering the solution of the Plateau problem as practically a corollary of the polyhedral mapping theorem of Koebe.

The exact statement of this mapping theorem is the following:

Let Π *denote any simply-connected polyhedral surface bounded by a polygon* P*; then there exists a one-one continuous map of* Π*, including its boundary, on the unit circle, including its circumference, which is conformal at every interior point of* Π *not a vertex.*

The existence proof proceeds, in outline. as follows: Let Γ denote the given contour and $m(\Gamma)$ the least possible limit that can be approached by the area of a polyhedral surface whose boundary tends to Γ. If Π_m denote a sequence of polyhedra with boundaries P_m tending to Γ and areas tending to $m(\Gamma)$, we then represent Π_m conformally on the disc of the unit circle, whereby a one-one continuous correspondence is set up between the circumference C and P_m. This induces in the limit as $m \to \infty$ a one-one continuous correspondence between C and Γ, which determines a definite parametric representation of Γ: $x_i = g_i(\theta)$, where θ denotes the polar angle on C. Then the required minimal surface is $x_i = H_i(u, v)$, where these are the harmonic functions in the interior of the unit circle whose boundary values are $g_i(\theta)$, as defined by Poisson's integral.

1. The functionals A(g) and S(g)

We shall deal with vector functions g of n components:

$$x_i = g_i(\theta), \qquad i = 1, 2, \cdots, n, \tag{1.1}$$

where θ denotes the polar angle on the circumference of the unit circle, and x_i are interpreted as the coordinates of a point in

n-dimensional euclidean space. It will be sufficient for our purpose to suppose that the vector function $g(\theta) = [g_1(\theta), g_2(\theta), \cdots, g_n(\theta)]$ is everywhere continuous or admits at most a denumerable infinity of discontinuities all of the first kind, i.e. where as θ approaches to a certain fixed point θ_0, distinct one-sided limits—the point P and the point Q, with $P \neq Q$—are approached by $g(\theta)$. We suppose too that the point $x_i = g_i(\theta)$ always remains in some given finite region of n-dimensional space:

$$|g_i(\theta)| \leqq M, \tag{1.2}$$

where M is fixed once and for all.

Let

$$x_i = H_i(u, v) \tag{1.3}$$

denote the harmonic functions in the interior of the unit circle which are determined according to Poisson's integral by the boundary functions (1.1); then denoting by $w = u + iv$ a complex variable ranging over the interior of the unit circle, $|w| < 1$, we may express these harmonic functions as

$$x_i = \Re F_i(w) \tag{1.4}$$

with

$$F_i(w) = \frac{1}{2\pi}\int_C \frac{e^{i\theta} + w}{e^{i\theta} - w} g_i(\theta)d\theta, \tag{1.5}$$

C denoting the circumference of the unit circle. The equations (1.3) or (1.4) represent a *harmonic surface H*. When the functions (1.1) are everywhere continuous, so representing a continuous closed curve Γ, the surface H is bounded by Γ, because by a well-known property of Poisson's integral the harmonic functions (1.3) then attach continuously to the boundary values (1.1). For a point θ_0 where a discontinuity of the first kind presents itself, the corresponding part of the boundary of H is the *line-segment PQ*; this results from the fact that the limiting value of $H_i(u, v)$ in the approach of (u, v) to the point θ_0 varies

linearly with the direction of approach. The boundary of H is then made up of the points in n-space which correspond to the points of continuity of $g(\theta)$, together with the line-segments PQ corresponding to the points of discontinuity.

The element of length of the harmonic surface H is

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2, \tag{1.6}$$

where

$$E = \sum_{i=1}^{n}\left(\frac{\partial H_i}{\partial u}\right)^2, \quad F = \sum_{i=1}^{n}\left(\frac{\partial H_i}{\partial u}\right)\left(\frac{\partial H_i}{\partial v}\right), \quad G = \sum_{i=1}^{n}\left(\frac{\partial H_i}{\partial v}\right)^2. \tag{1.7}$$

Consider now the two double integrals:

$$A(g) = \int\int_D \tfrac{1}{2}(E + G)d\sigma = \tfrac{1}{2}\sum_{i=1}^{n}\left[\left(\frac{\partial H_i}{\partial u}\right)^2 + \left(\frac{\partial H_i}{\partial v}\right)^2\right]d\sigma \tag{1.8}$$

and

$$S(g) = \int\int_D \sqrt{EG - F^2}\, d\sigma, \tag{1.9}$$

where the domain of integration D is the interior of the unit circle, of which $d\sigma$ represents the element of area. As the notation indicates, these are functionals of g, being completely determined as soon as g is given. $A(g)$ is one-half the sum of the Dirichlet functionals formed for each coordinate of H; $S(g)$ is the area of the surface H. Since the integrands may not remain bounded in D, these functionals must be defined more exactly as the respective limits for $\rho \to 1$ of

$$A_\rho(g) = \int\int_{D_\rho} \tfrac{1}{2}(E + G)d\sigma \tag{1.8'}$$

and

$$S_\rho(g) = \int\int_{D_\rho} \sqrt{EG - F^2}\, d\sigma, \tag{1.9'}$$

where D_ρ denotes the interior of the circle of radius $\rho < 1$ about the origin as center. Since in both cases the element of integration is positive, it follows that $A_\rho(g)$ and $S_\rho(g)$ must increase with ρ, and therefore the corresponding limits $A(g)$ and $S(g)$ must exist, finite or positively infinite.

By applying Green's theorem to the double integral (1.8′) there was derived in the cited paper the following formula giving $A(g)$ explicitly in terms of g:

$$A(g) = \frac{1}{4\pi}\int_C\int_C \frac{\sum_{i=1}^{n}[g_i(\theta) - g_i(\varphi)]^2}{4\sin^2\frac{\theta-\varphi}{2}}\,d\theta d\varphi; \tag{1.10}$$

but we shall carry through the work of the present paper without the use of this formula.

2. Preliminary theorems

In order that our main argument, as developed in § 4, may proceed without interruption, we gather together in this section five simple preliminary theorems which will be appealed to in the later discussion. These theorems and their proofs are, almost entirely, a reproduction of certain parts of the paper cited under [1].

THEOREM I.

$$S(g) \leqq A(g), \tag{2.1}$$

and the equality sign holds when and only when

$$\sum_{i=1}^{n} F_i'^2(w) = 0. \tag{2.2}$$

Here the functions $F_i(w)$ are those of (1.4), whose real parts are the coordinates of the harmonic surface H, and the condition (2.2) expresses, according to the formulas of Weierstrass, that the surface H is *minimal*.

Proof. The difference of the squares of the integrands of $A(g)$ and $S(g)$ is

$$\tfrac{1}{4}(E+G)^2 - (EG - F^2) = \tfrac{1}{4}(E-G)^2 + F^2 \geqq 0.$$

Hence

$$\tfrac{1}{2}(E+G) \geqq \sqrt{EG - F^2},$$

and consequently

$$A(g) \geqq S(g),$$

where on account of the continuity (indeed, analyticity) of E, F, G the equality sign holds when and only when

$$E - G \equiv 0, \qquad F \equiv 0. \tag{2.3}$$

Now from

$$H_i(u, v) = \Re F_i(w)$$

we have

$$F_i'(w) = \frac{\partial H_i}{\partial u} - i\frac{\partial H_i}{\partial v}, \tag{2.4}$$

and therefore with reference to (1.7)

$$\sum_{i=1}^{n} F_i'^2(w) = (E - G) - 2iF,$$

so that the conditions (2.3) are equivalent to

$$\sum_{i=1}^{n} F_i'^2(w) \equiv 0.$$

THEOREM II. *$A(g)$ is lower semi-continuous.*

By this we mean: consider any sequence of vector functions

$$g^{(1)}(\theta), \quad g^{(2)}(\theta), \quad \cdots, \quad g^{(m)}(\theta), \quad \cdots \tag{2.5}$$

which remain uniformly bounded according to (1.2) and tend to a limit:

$$\lim_{m \to \infty} g^{(m)}(\theta) = g(\theta). \tag{2.6}$$

The approach need not be uniform and the possible denumerable infinity of points of discontinuity of the functions $g^{(m)}(\theta)$ and $g(\theta)$ may be disregarded in this limit relation. Then will

$$A(g) \leqq \liminf_{m \to \infty} A(g^{(m)}). \tag{2.7}$$

In other words, if

$$m = m_1, \quad m_2, \quad \cdots, \quad m_k, \quad \cdots$$

represent any subsequence of (2.5) for which $A(g^{(m)})$ tends to a limit L, finite or infinite (such subsequences, of course, always exist), then

$$A(g) \leqq L. \tag{2.8}$$

In case $A(g) = +\infty$, this lower semi-continuity amounts to continuity; we then have

$$\lim_{m \to \infty} A(g^{(m)}) = +\infty \tag{2.9}$$

for every sequence $g^{(m)}$ tending to g.

Proof. The proof of Theorem II rests on the following simple lemma, well-known for functions of n real variables,[3] and extended by the writer [4] to functionals on any Fréchet L-set; the proof of the lemma is evident and will not be repeated here.

LEMMA. *If a functional $A(g)$ can be expressed as the limit of a continuous functional $A_\rho(g)$ which tends to $A(g)$ in increasing, then $A(g)$ is lower semi-continuous.*

That this lemma applies to the functional $A(g)$ defined by (1.8) with the functional $A_\rho(g)$ defined by (1.8′) is easy to see. First, $A_\rho(g)$ is a continuous functional of g for every fixed $\rho < 1$, since its integrand remains uniformly bounded, and under that condition one may pass to the limit under the sign of integration:

$$\lim_{m \to \infty} A_\rho(g^{(m)}) = A_\rho(g). \tag{2.10}$$

To see the uniform boundedness, we observe that by (2.4) and (1.7) we may write (1.8′)

$$A_\rho(g) = \int\int_{D_\rho} \tfrac{1}{2} \sum_{i=1}^{n} |F_i'(w)|^2 d\sigma. \tag{2.11}$$

By differentiation under the integral sign of (1.5),

$$F_i'(w) = \frac{1}{\pi} \int_C \frac{e^{i\theta}}{(e^{i\theta} - w)^2} g_i(\theta) d\theta, \tag{2.12}$$

and consequently under the condition $|w| \leqq \rho$ we have with attention to (1.2),

$$|F_i'(w)| \leqq \frac{2M}{(1-\rho)^2}, \tag{2.13}$$

[3] See Carathéodory, *Vorlesungen über reele Funktionen*, Leipzig, 1918, p. 175.

[4] Loc. cit., p. 282.

from which results immediately the uniform boundedness of the integrand of (2.11).

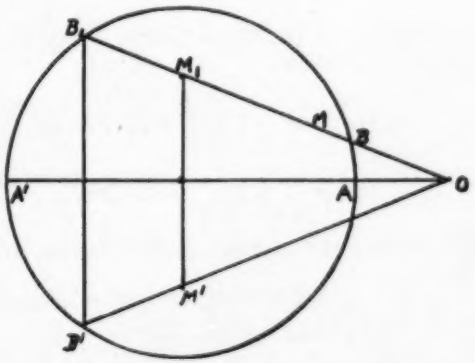
That $A_\rho(g)$ tends to $A(g)$ in increasing is immediate from the observation that the element of integration in $A(g)$ is always positive.

THEOREM III. *If $g(\theta)$ has a point of discontinuity with unequal one-sided limits, then $A(g) = +\infty$.*

Proof. There is no loss of generality in supposing the discontinuity to be located at the point $z = 1$, since this may be accomplished by a rotation of the unit circle, which changes nothing essential. We will suppose the one-sided limits of $g(\theta)$ at $z = 1$, or A, to be the points P and Q of n-dimensional space with the coordinates a_i and b_i respectively. By assumption, the distance l between these two points, with

$$l^2 = \sum_{i=1}^{n} (a_i - b_i)^2, \tag{2.14}$$

is greater than zero.



We now define a conformal transformation of the unit circle into itself as follows. Let O be any point on the diameter $A'A$ prolonged. If M is any point in the interior of the unit circle,

we first define a transformation $M \to M_1$ by the condition that on the ray $OBMM_1B_1$

$$OM \cdot OM_1 = OB \cdot OB_1.$$

We then combine $M \to M_1$ with a reflection $M_1 \to M'$ in the diameter $A'A$ to produce the transformation

$$\mathfrak{T} = M \to M'.$$

This is evidently conformal, converts the unit circle into itself in a one-one continuous way. circumference into circumference and interior into interior, and has the linear fractional form

$$w' = \frac{aw + b}{cw + d}. \tag{2.15}$$

The transformation (2.15) acts on the boundary functions $g_i(\theta)$ and the functions $F_i(w)$ determined by them to produce boundary functions $h_i(\theta)$ determining functions $G_i(w)$, where

$$G_i\left(\frac{aw + b}{cw + d}\right) = F_i(w). \tag{2.16}$$

It is evident that the functional $A(g)$ is invariant under this transformation. For writing

$$A(g) = \int\int_D \tfrac{1}{2} \sum_{i=1}^{n} |F_i'(w)|^2 d\sigma,$$

$$A(h) = \int\int_D \tfrac{1}{2} \sum_{i=1}^{n} |G_i'(w)|^2 d\sigma,$$

we have, if λ denote the modulus of the conformal transformation:[5]

$$\lambda = \left|\frac{dw'}{dw}\right| = \left|\frac{ad - bc}{(cw + d)^2}\right|,$$

that

$$d\sigma' = \lambda d\sigma, \quad \text{while} \quad |G_i'(w)| = \frac{1}{\lambda} |F_i'(w)|;$$

[5] Evidently λ is nowhere zero or infinite in the interior of the unit circle.

consequently

$$A(h) = A(g). \tag{2.17}$$

Imagine now that O takes a sequence of positions

$$O_1, \quad O_2, \quad \cdots, \quad O_m, \quad \cdots \tag{2.18}$$

all on $A'A$ produced and tending to A as limit. It is obvious that if B' is any fixed point on the unit circle, then the image of B' by $\mathfrak{T}^{-1}$ tends to A from above or below according as B' is on the lower or upper semi-circumference. Hence $h_i(\theta)$ tends to the function

$$k_i(\theta) = \begin{cases} a_i & \text{for} \quad 0 < \theta < \pi \\ b_i & \text{for} \quad \pi < \theta < 2\pi. \end{cases}$$

The function of w determined by the boundary function $k_i(\theta)$ is

$$\begin{aligned} H_i(w) &= \frac{1}{2\pi}\int_0^{\pi} \frac{e^{i\theta} + w}{e^{i\theta} - w} a_i d\theta + \frac{1}{2\pi}\int_{\pi}^{2\pi} \frac{e^{i\theta} + w}{e^{i\theta} - w} b_i d\theta \\ &= \frac{a_i}{\pi i}\log\,(e^{i\theta/2} - we^{-(i\theta/2)})\Big]_0^{\pi} + \frac{b_i}{\pi i}\log\,(e^{i\theta/2} - we^{-(i\theta/2)})\Big]_{\pi}^{2\pi} \\ &= \frac{a_i}{\pi i}\log\, i\frac{1+w}{1-w} + \frac{b_i}{\pi i}\log\, i\frac{1-w}{1+w} \\ &= \frac{a_i - b_i}{\pi i}\log\, i\frac{1+w}{1-w}. \end{aligned}$$

From this we calculate

$$H_i'(w) = \frac{2(a_i - b_i)}{\pi i} \cdot \frac{1}{1 - w^2}, \tag{2.19}$$

and consequently, with reference to (2.14),

$$\tfrac{1}{2}\sum_{i=1}^{n} |H_i'(w)|^2 = \frac{2l^2}{\pi^2} \cdot \frac{1}{|1 - w^2|^2}.$$

Hence

$$A(k) = \frac{2l^2}{\pi^2}\int\int_D \frac{1}{|1 - w^2|^2} d\sigma,$$

and we can easily see that the value of this double integral is $+\infty$:

$$A(k) = +\infty. \tag{2.20}$$

For if we introduce polar coordinates with the point $w = 1$ as pole, then the element of integration is, for $w \to 1$, asymptotically equal to

$$\frac{1}{4\rho^2} \cdot \rho d\rho d\theta = \frac{1}{4}\frac{d\rho}{\rho} d\theta.$$

The integral of this over the area

$$\rho_1 < \rho < \rho_0, \qquad -\alpha < \theta < \alpha \qquad \left(0 < \alpha < \frac{\pi}{2}\right)$$

is

$$\frac{\alpha}{2} \log \frac{\rho_0}{\rho_1},$$

which becomes $+\infty$ when $\rho_1 \to 0$. Consequently $+\infty$ is the value of $A(k)$ since the part of the unit circle outside the sector of angle 2α and radius ρ_0 with vertex at $w = 1$ must make a non-negative contribution to the value of the double integral.

Let now

$$h^{(1)}(\theta),\ h^{(2)}(\theta),\ \cdots,\ h^{(m)}(\theta),\ \cdots$$

be the boundary functions derived from $g(\theta)$ by the transformations

$$\mathfrak{T}_1,\ \mathfrak{T}_2,\ \cdots,\ \mathfrak{T}_m,\ \cdots$$

associated with the points (2.18); then by the invariance property (2.17),

$$A(h^{(1)}) = A(h^{(2)}) = \cdots = A(h^{(m)}) = \cdots = A(g),$$

accordingly

$$\lim_{m \to \infty} A(h^{(m)}) = A(g). \tag{2.21}$$

By the lower semi-continuity of $A(g)$, according to (2.8), it follows from (2.21) and

$$\lim_{m \to \infty} h^{(m)}(\theta) = k(\theta)$$

that

$$A(k) \leqq \nu A(g);$$

consequently by (2.20),

$$A(g) = +\infty,$$

which was to be proved.

THEOREM IV. *If $g(\theta)$ has a point of discontinuity with unequal one-sided limits, then*

$$\sum_{i=1}^{n} F_i'^2(w) \not\equiv 0.$$

Proof. Using the same figure and sequence of linear fractional transformations $\mathfrak{T}_m$, as in the proof of the preceding theorem, we observe first from (2.16) the relation

$$(2.22) \qquad \frac{ad - bc}{(cw + d)^2} G_i^{(m)'}\left(\frac{aw + b}{cw + d}\right) = F_i'(w), \quad m = 1, 2, 3, \cdots.$$

We desire to derive contradiction from the hypothesis

$$\sum_{i=1}^{n} F_i'^2(w) \equiv 0.$$

If would follow from this according to (2.22) that

$$(2.23) \qquad \sum_{i=1}^{n} [G_i^{(m)'}(w)]^2 \equiv 0, \quad m = 1, 2, 3, \cdots.$$

Now by (2.12),

$$G_i^{(m)'}(w) = \frac{1}{\pi} \int_C \frac{e^{i\theta}}{(e^{i\theta} - w)^2} h_i^{(m)}(\theta) d\theta,$$

$$H_i'(w) = \frac{1}{\pi} \int_C \frac{e^{i\theta}}{(e^{i\theta} - w)^2} k_i(\theta) d\theta,$$

the notation being that of the proof of Theorem III. For any fixed value of w, the first integrand evidently remains uniformly bounded as it approaches to the second (see (1.2)), and hence

$$\lim_{m \to \infty} G_i^{(m)'}(w) = H_i'(w).$$

Consequently

$$\lim_{n\to\infty}\sum_{i=1}^{n}[G_i^{(n)\prime}(w)]^2=\sum_{i=1}^{n}H_i'^2(w)=-\frac{4l^2}{\pi^2}\cdot\frac{1}{(1-w^2)^2}$$

according to (2.19). But since $l\neq 0$, this is in contradiction with (2.23).

THEOREM V. *If $g(\theta)$ is constant on an arc of C, but not on all of C, then*

$$\sum_{i=1}^{n}F_i'^2(w)\not\equiv 0.$$

Proof. By (2.12),

$$wF_i'(w)=\frac{1}{2\pi}\int_C\frac{2we^{i\theta}}{(e^{i\theta}-w)^2}g_i(\theta)d\theta.$$

Writing $w=\rho e^{i\alpha}$ and taking the imaginary part of each side, we get

$$\Im wF_i'(w)=-\frac{1}{2\pi}\int_C\frac{2\rho(1-\rho^2)\sin(\theta-\alpha)}{[1-2\rho\cos(\theta-\alpha)+\rho^2]^2}g_i(\theta)d\theta.$$

Except for the minus sign, the last expression is identical with one studied by Fatou in his thesis.[6] By a simple integration by parts, Fatou shows that at every point where the derivative $g_i'(\theta)$ exists and is continuous the expression in question has the unique limit $g_i'(\theta)$ when (ρ,α) approaches to $(1,\theta)$ in an arbitrary manner.

In the present case $g_i(\theta)$ is supposed to be constant on a certain arc AB of C; therefore $g_i'(\theta)$ has the continuous value zero on this arc; hence for θ any interior point of AB,

$$\lim_{w\to e^{i\theta}}\Im wF_i'(w)=0,\quad i=1,2,\cdots,n. \tag{2.24}$$

Suppose, contrary to what we wish to prove, that

$$\sum_{i=1}^{n}F_i'^2(w)\equiv 0.$$

[6] P. Fatou, "Séries trigonométriques et séries de Taylor," *Acta Mathematica*, vol. 30 (1906), pp. 347–348.

Then

$$\sum_{i=1}^{n} w^2 F_i'^2(w) = \sum_{i=1}^{n} [\Re w F_i'(w) + i\Im w F_i'(w)]^2 \equiv 0,$$

and taking real parts, we have

$$\sum_{i=1}^{n} [\Re w F_i'(w)]^2 \equiv \sum_{i=1}^{n} [\Im w F_i'(w)]^2.$$

From this together with (2.24) it follows that

$$\lim_{w \to e^{i\theta}} \Re w F_i'(w) = 0, \quad i = 1, 2, \cdots, n, \tag{2.25}$$

or, combining (2.25) and (2.24), that

$$\lim_{w \to e^{i\theta}} w F_i'(w) = 0, \quad i = 1, 2, \cdots, n,$$

for θ any interior point of AB.

Since, therefore, the limiting value of $wF_i'(w)$ when w approaches to any point of the circular arc AB is real, we may apply the symmetry principle of Riemann-Schwarz to prolong the function analytically across this arc, attaching conjugate imaginary values to points inverse with respect to the arc. The function $wF_i'(w)$ will then be equal to zero on an arc interior to a domain of regularity, and is therefore identically zero. Hence $F_i(w)$ is identically equal to a constant. Since $\Re F_i(w)$ reduces to $g_i(\theta)$ on the circumference C, we thus have contradiction with that part of the hypothesis of our theorem which stipulates that $g(\theta)$ shall not be constant on all of C.

3. Topological automorphisms of the unit circle [7]

Let Γ denote any Jordan curve in euclidean space of n-dimensions; it is representable by equations

$$x_i = f_i(t), \quad i = 1, 2, \cdots, n, \tag{3.1}$$

where $f_i(t)$ are continuous functions of the parameter t, which we interpret as polar angle on the unit circle. From t we pass to any other parameter θ by a one-one continuous transformation

[7] See § 3 and § 8 of the cited paper.

of the unit circle into itself,

$$t = \omega(\theta), \tag{3.2}$$

the equations (3.1) becoming by the substitution (3.2)

$$x_i = f_i(\omega(\theta)) = g_i(\theta). \tag{3.3}$$

Of basic importance is the remark that *the totality of topological transformations $t = \omega(\theta)$ of the unit circle into itself forms a compact set.* This means that out of any infinite set of such transformations a subsequence can be selected which tends to a limit, itself a topological automorphism of the unit circle, either proper or improper (in a certain sense which will soon be specified).

This fact is best seen by means of the following geometrical representation on a torus of the topological automorphisms of the unit circle. We interpret θ as angle along a parallel and t as angle along a meridian of the torus; then a one-one continuous relation between θ and t is represented by a continuous closed curve on the torus which is intersected once and only once by each meridian and by each parallel: such a curve we may call *cyclically monotonic.* As limit form of a cyclically monotonic curve we may have one which contains segments of a parallel or segments of a meridian, and even, as extreme limit form, one which consists of a parallel together with a meridian. We include these forms as defining " improper " topological transformations of the unit circle into itself, and the extreme limit form as defining a " degenerate " transformation. In an improper parametric representation of Γ, an arc of Γ less than all of Γ corresponds to a single point of C, or an arc of C less than all of C corresponds to a single point of Γ. In a degenerate parametric representation, a single point of Γ corresponds to all of C, and a single point of C to all of Γ.

The totality of cyclically monotonic curves on the torus, proper, improper and degenerate, forms a compact closed set. That this is the fact results immediately from the following theorem:[8]

[8] M. Fréchet, "Sur quelques points du calcul fonctionel," *Rendiconti Palermo*, vol. 22 (1906), p. 65.

An infinite set of curves is compact if the curves are all contained in a finite region and have uniformly bounded lengths.

In the present case the finite region is the torus, and the uniform bound for the length of all the curves is the length of a parallel plus the length of a meridian; this is to be seen immediately from the cyclically monotonic character of the curves. The closure of the set of curves (the limit of any infinite sequence of curves of the set belongs to the set) is obvious for the same reason.

We shall find it convenient to fix three distinct points θ_1, θ_2, θ_3 of the unit circle and three distinct points t_1, t_2, t_3, and to consider only those topological transformations which convert θ_1, θ_2, θ_3 into t_1, t_2, t_3. The corresponding curves on the torus pass through the three fixed points (θ_1, t_1), (θ_2, t_2), (θ_3, t_3), no two of which are on the same parallel or the same meridian. They define a subset of topological automorphisms of the unit circle which is compact and closed, and which, while containing improper representations, has the advantage of not containing any degenerate ones. No generality is lost by this fixation of three points, since everything is invariant under a conformal, or linear fractional, transformation of the unit circle into itself, which involves three arbitrary constants.

4. Existence of a minimal surface bounded by a Jordan curve Γ in the case $m(\Gamma)$ finite

These preliminaries established, we now proceed to the existence theorem whose proof is the main purpose of the present paper.

With any Jordan curve Γ we may associate a definite positive number $m(\Gamma)$, finite or infinite, introduced by Lebesgue[9] and called by him the minimum area of Γ.

Definition. $m(\Gamma)$ is the least limit which can be approached by the area of a polyhedral surface Π whose polygonal boundary P tends to Γ.

[9] H. Lebesgue, "Intégrale, longueur, aire," *Annai di Matematica*, (3), vol. 7 (1902), p. 304.

Assumption. In this section we restrict ourselves to the case where $m(\Gamma)$ is finite.

$m(\Gamma)$ is, for instance, finite whenever Γ is a polygon or any rectifiable contour. The proof is evident.

Let Σ denote any continuous surface bounded by Γ; then, according to the Lebesgue definition of area, the area $S(\Sigma)$ of Σ is the inferior limit of the areas of polyhedral surfaces which tend to Σ. Since the boundaries of such polyhedral surfaces must tend to Γ, it is obvious from the definition of $m(\Gamma)$ that

$$m(\Gamma) \leqq S(\Sigma). \tag{4.1}$$

This remark made, we observe next that by definition of $m(\Gamma)$ there exists a sequence of polyhedral surfaces

$$\Pi_m, \quad m = 1, 2, 3, \cdots,$$

whose polygonal boundaries

$$P_m, \quad m = 1, 2, 3, \cdots$$

tend to Γ, and such that the areas of the polyhedral surfaces tend to $m(\Gamma)$:

$$\lim_{m \to \infty} S(\Pi_m) = m(\Gamma). \tag{4.2}$$

The polygons P_m are representable by equations

$$x_i = f_i^{(m)}(t), \tag{4.3}$$

where the functions $f_i^{(m)}(t)$ tend uniformly to those defining Γ:

$$x_i = f_i(t). \tag{4.4}$$

We now apply Koebe's theorem, as formulated in the introduction. There exists a conformal map of Π_m on the interior of the unit circle which remains one-one and continuous (even piece-wise analytic) as between P_m and the circumference C. In this way a one-one continuous relation is established between t the parameter on P_m and θ the angle on C:

$$t = \omega_m(\theta). \tag{4.5}$$

By adjoining a conformal transformation of the unit circle into itself, which involves three arbitrary constants, we can arrange that three fixed points θ_1, θ_2, θ_3 shall correspond respectively to three fixed points t_1, t_2, t_3, the fixed points being the same for all the polygons P_m.

From the infinite sequence of topological transformations (4.5) we can, according to § 3, select a subsequence

$$m = m_1, m_2, \cdots, m_k, \cdots, \tag{4.6}$$

which tends to a limit

$$t = \omega(\theta) \tag{4.7}$$

that may be a proper or improper topological automorphism of the unit circle, but not a degenerate one (on account of the three fixed points condition, see end § 3).

Keeping only the values of m belonging to the sequence (4.6), we then make the transformation of parameter (4.5) in the equations (4.3) of each polygon P_m, so obtaining the parametric equations

$$x_i = f_i^{(m)}(\omega_m(\theta)) = g_i^{(m)}(\theta) \tag{4.8}$$

for these polygons. If we make the limit transformation (4.7) in the equations (4.4) of Γ, we get

$$x_i = f_i(\omega(\theta)) = g_i(\theta). \tag{4.9}$$

As far as we know at this stage, this parametric representation of Γ might be improper in that the functions $g_i(\theta)$ might present (at most) a denumerable infinity of discontinuities of the first kind, or they might be constant on certain arcs of C. We shall prove that in fact (4.9) is a proper parametric representation of Γ, and that the harmonic surface determined by it is the minimal surface whose existence we are trying to establish.

First, it is evident that

$$\lim_{m \to \infty} g_i^{(m)}(\theta) = g_i(\theta),$$

abstraction being made of any values of θ for which $g_i(\theta)$ may be discontinuous. These are at most denumerably infinite in number, form thus a set of zero measure, and therefore do not affect any limit relations involving integrals which depend on $g^{(m)}(\theta)$ and $g(\theta)$.

The conformal map of Π_m on the unit circular disc gives a parametric representation of Π_m,

$$x_i = x_i(u, v), \tag{4.10}$$

where the functions $x_i(u, v)$ are piece-wise analytic and such that

$$E \equiv G, \qquad F \equiv 0, \tag{4.11}$$

where E, F, G are the coefficients in the first fundamental form of the surface Π_m:

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2.$$

Let us denote by $D(\Pi_m)$ one-half the sum of the n Dirichlet functionals formed for the coordinates of Π_m:

$$D(\Pi_m) = \tfrac{1}{2}\sum_{i=1}^{n}\int\int_D\left[\left(\frac{\partial x_i}{\partial u}\right)^2 + \left(\frac{\partial x_i}{\partial v}\right)^2\right]d\sigma = \int\int_D \tfrac{1}{2}(E+G)d\sigma.^{10}$$

Comparing with

$$S(\Pi_m) = \int\int_D \sqrt{EG - F^2}\, d\sigma,$$

we have immediately from (4.11):

$$S(\Pi_m) = D(\Pi_m). \tag{4.12}$$

Consider now $A(g^{(m)})$; if $H_i^{(m)}(u, v)$ is the harmonic function determined by the boundary values $g_i^{(m)}(\theta)$, we have by definition

$$A(g^{(m)}) = \tfrac{1}{2}\sum_{i=1}^{n}\int\int_D\left[\left(\frac{\partial H_i^{(m)}}{\partial u}\right)^2 + \left(\frac{\partial H_i^{(m)}}{\partial v}\right)^2\right]d\sigma.$$

[10] If δ_j denote the curvilinear triangles on the unit circular disc which correspond to the various faces of the polyhedron, this integral means the sum of the integrals over δ_j, and the integral over δ_j means the upper bound of the integral over all regions (polygons would suffice) which with their boundaries are interior to δ_j.

Now the harmonic function $H_i^{(m)}(u, v)$ has the same boundary values as the function $x_i(u, v)$ of (4.10); hence by the property possessed by harmonic functions of minimising Dirichlet's integral for given boundary values,[11] we have

$$A(g^{(m)}) \leqq D(\Pi_m),$$

consequently by (4.12)

$$A(g^{(m)}) \leqq S(\Pi_m). \tag{4.13}$$

By (4.2), $S(\Pi_m)$ tends to the finite quantity $m(\Gamma)$ when $m = 1, 2, 3, \cdots$; consequently it still does so when m varies through the subsequence (4.6); hence by (4.13) $A(g^{(m)})$ remains bounded; it follows that there exists a subsequence of (4.6)

$$m = m_1', m_2', \cdots, m_k', \cdots \tag{4.14}$$

such that

$$\lim_{m \to \infty} A(g^{(m)}) = L, \tag{4.15}$$

where necessarily, by (4.13) and (4.2),

$$L \leqq m(\Gamma). \tag{4.16}$$

By the lower semi-continuity of $A(g)$, as expressed in (2.8) it follows from (4.15) that

$$A(g) \leqq L,$$

therefore by (4.16)

$$A(g) \leqq m(\Gamma). \tag{4.17}$$

The value of $A(g)$ being thus finite, it follows at once that the parametric representation g of Γ cannot be improper by the presence of discontinuities with unequal one-sided limits, for

[11] See Hurwitz-Courant, Funktionentheorie, Berlin, 1922, p. 335, Hilfssatz II. The comparison functions are all functions continuous in and on the unit circle, having the given boundary values, and such that the unit circle can be divided into a finite number of curvilinear triangles in the interior of each of which the function is continuously differentiable: the functions $x_i(u, v)$ are thus eligible for comparison.

then we would have according to Theorem III: $A(g) = +\infty$.

Let H denote the harmonic surface determined by the parametric representation g of Γ. The only way in which g might still be improper would be for the transformation of parameter (4.7) to make an arc of the unit circle C correspond to a point of Γ. But this circumstance. even if it occurred, could not prevent us from regarding H as a surface bounded by Γ in the exact sense of the following definition.

Definition. A continuous surface Σ bounded by Γ is the image of the unit circular disc by a continuous transformation which makes the circumference C correspond in a one-one way to Γ.

For consider the inverse transformation to (4.7):

$$\theta = \omega^{-1}(t). \tag{4.18}$$

In case this is improperly monotonic, it can be represented as the limit for $\rho \to 1$ of properly monotonic transformations

$$\theta = \omega^{-1}(t, \rho), \quad 0 < \rho < 1. \tag{4.19}$$

If we make the transformation of polar angle (4.18) on the circumference C and the transformation (4.19) on every concentric circle of radius $\rho < 1$, then we derive a representation of H in terms of parameters (ρ, t) whereby the parametric equations of Γ become the original ones, $x_i = f_i(t)$, representing Γ as a one-one continuous image of C. We are thus justified in considering H as, in any event, bounded by Γ.

Since H is then a particular surface bounded by Γ. we have for its area $S(g)$, according to (4.1):

$$m(\Gamma) \leqq S(g). \tag{4.20}$$

On the other hand, we have by Theorem I

$$S(g) \leqq A(g). \tag{4.21}$$

Now, combining (4.21), (4.17) and (4.20), we form the chain of inequalities

$$S(g) \leqq A(g) \leqq m(\Gamma) \leqq S(g). \tag{4.22}$$

On account of the identity of the first and last terms, the equality must hold throughout:

$$S(g) = A(g) = m(\Gamma). \tag{4.23}$$

This equation contains all the desired results.

1°. From $S(g) = A(g)$ it follows by Theorem I that the surface H is minimal in the sense of the Weierstrass condition $\sum_{i=1}^{n} F_i'^2(w) \equiv 0$.

2°. The relation $S(g) = m(\Gamma)$ expresses, in conjunction with (4.1) that H has the least area of any surface bounded by Γ.

3°. We have already ruled out the possibility of the parametric representation g of Γ being improper by the presence of a discontinuity. The possibility of g being improper by being constant on an arc of C is now ruled out by Theorem V. The minimal surface H is therefore bounded by Γ.

5. The case $m(\Gamma) = +\infty$ [12]

The significance of the condition $m(\Gamma) = +\infty$ is that it is impossible to span any continuous surface of finite area within the given contour. In this case, the existence of a minimal surface M bounded by Γ can be proved by means of a simple limit process in which Γ is regarded as a limit of polygons. We have proved too [13] that although the area of M is $+\infty$, it still has the least area property in the sense that every contour Γ' drawn upon M and having no point in common with the boundary Γ intercepts upon M a finite area $\leqq$ any other area bounded by Γ'. We shall not reproduce this proof here, but content ourselves with establishing the existence of a minimal surface in the sense of the Weierstrass formulas.

[12] This section is essentially a reproduction of §§ 19 and 20 of the cited paper.

[13] "The Least Area Property of the Minimal Surface Determined by an Arbitrary Jordan Contour," Proceedings of the National Academy of Sciences, vol. 17, no. 4 (April, 1931), pp. 211–216.

The Jordan curve Γ can be expressed as the limit of a sequence of non-self-intersecting polygons

$$\Gamma^{(1)}, \Gamma^{(2)}, \cdots, \Gamma^{(m)}, \cdots.$$

These may be represented parametrically by equations

$$x_i = f_i^{(m)}(t) \tag{5.1}$$

where the functions $f_i^{(m)}(t)$ tend uniformly to

$$x_i = f_i(t), \tag{5.2}$$

some fixed parametric representation of Γ.

For each polygon $\Gamma^{(m)}$, the minimum area $m(\Gamma^{(m)})$ is evidently finite, and hence by § 4 there exists a parametric representation of $\Gamma^{(m)}$,

$$x_i = g_i^{(m)}(\theta), \tag{5.3}$$

such that the harmonic surface determined by it:

$$x_1 = \Re F_i^{(m)}(w), \tag{5.4}$$

$$F_i^{(m)}(w) = \frac{1}{2\pi}\int_C \frac{e^{i\theta} + w}{e^{i\theta} - w} g_i^{(m)}(\theta)d\theta,$$

is minimal:

$$\sum_{i=1}^{n} [F^{(m)\prime}(w)]^2 \equiv 0. \tag{5.5}$$

The parameter θ on $\Gamma^{(m)}$ is derived from the parameter t by a topological automorphism of the unit circle,

$$t = \omega_m(\theta). \tag{5.6}$$

We have seen that no generality is lost by supposing that in (5.6) certain three distinct fixed points $\theta_1, \theta_2, \theta_3$ correspond to certain three distinct points t_1, t_2, t_3.

According to § 3, we can select from (5.6) a subsequence:

$$m = m_1, m_2, \cdots, m_k, \cdots \tag{5.7}$$

tending to a limit

$$t = \omega(\theta), \tag{5.8}$$

which may be a proper or improper topological automorphism of the unit circle, but not a degenerate one, on account of the condition of the three distinct fixed points.

Let us make the transformation (5.8) in the parametric equations (5.2) of Γ:

$$x_i = f_i(\omega(\theta)) = g_i(\theta). \tag{5.9}$$

Then the harmonic surface determined by the representation g of Γ:

$$x_i = \Re F_i(w), \tag{5.10}$$

$$F_i(w) = \frac{1}{2\pi}\int_C \frac{e^{i\theta} + w}{e^{i\theta} - w}\, g_i(\theta)d\theta,$$

is minimal:

$$\sum_{i=1}^{n} F_i'^2(w) \equiv 0. \tag{5.11}$$

For we have from (5.4) and (5.10) by differentiation under the integral sign:

$$F^{(m)\prime}(w) = \frac{1}{\pi}\int \frac{e^{i\theta}}{(e^{i\theta} - w)^2}\, g_i^{(m)}(\theta)d\theta, \tag{5.12}$$

$$F'(w) = \frac{1}{\pi}\int_C \frac{e^{i\theta}}{(e^{i\theta} - w)^2}\, g_i(\theta)d\theta. \tag{5.13}$$

For any fixed value of w with $|w| < 1$, the integrand in (5.12) stays uniformly bounded for all m and θ; consequently we may pass to the limit under the sign of integration:

$$\lim_{m\to\infty} F^{(m)\prime}(w) = F'(w). \tag{5.14}$$

Hence (5.11) follows from the fact that (5.5) is obeyed for all values of m.

Before we can say that the minimal surface (5.10) is bounded by Γ, we have still to prove that the parametric representation g of Γ is proper. But this is an immediate consequence of Theorems IV and V. In view of the condition (5.11), Theorem IV shows immediately that g can have no discontinuities, and Theorem V that g cannot be constant on any arc of C (the possibility of g being constant on all of C having already been ruled out by the condition of three distinct fixed points θ_1, θ_2, θ_3 of C corresponding to three distinct fixed points t_1, t_2, t_3 of Γ).

The existence of a minimal surface bounded by the arbitrary Jordan curve Γ is thus established.

STIRLING EXPANSIONS DERIVED BY MEANS OF FINITE DE LA VALLÉE-POUSSIN SUMMATION [1]

By Raymond D. Douglass

1. Introduction. Given an integral function $f(x)$ which is real-valued for x real, there exists a polynomial with real coefficients which agrees in value with $f(x)$ for $x = -nh, \cdots, -h, 0, h, \cdots, nh$ and which is uniquely determined by the condition that its degree shall not exceed $2n$. The sequence of polynomials obtained by allowing n to take the values 1, 2, 3, $\cdots$, is expressed by the Gauss interpolation series or by the Stirling interpolation series. These two expressions for the sequence of polynomials differ only in form, and we confine our attention to the latter. We shall be concerned, as will presently appear, with certain relations between the coefficients of Stirling's series and the de la Vallée-Poussin process of summation of divergent series, as developed in a recent paper by Rutledge.[2]

If $f(x)$ is a given integral function we have

$$f(x) = f(0) + G(x) + F(x), \tag{1}$$

where

$$G(x) = \frac{f(x) - f(-x)}{2}, \qquad F(x) = \frac{f(x) + f(-x) - 2f(0)}{2}. \tag{2}$$

Since we may write

$$f(x) = f(0) + G(x) + x \cdot \frac{F(x)}{x}, \tag{3}$$

[1] This paper was prepared under the supervision of Professor George Rutledge as a thesis for the degree of Doctor of Philosophy. The author is indebted to Professor Rutledge for his many valuable suggestions.

[2] Rutledge, "A Necessary and Sufficient Condition for Convergence of a Sequence of Lagrange Polynomials," *Jour. of Math. and Phys.*, vol. 9 (1930), pp. 261–267.

we see that to obtain the Stirling series for $f(x)$ we need merely the Stirling series for each of two odd functions. We shall, therefore, be concerned only with an odd integral function, to be denoted in general by $G(x)$, and with the associated Stirling series. Without loss of generality we take $h = 1$.

The Stirling series in question is then

$$\sum_{s=1}^{n} \delta_s x(x^2 - 1^2)(x^2 - 2^2) \cdots (x^2 - \overline{s-1}^2), \tag{4}$$

with $n = 1, 2, 3, \cdots$, and with $\delta_s = \frac{1}{(2s-1)!}\mu\Delta^{(2s-1)}$, the notation being the conventional one for mean central differences of odd order of the values

$$\cdots, G(-2), G(-1), 0, G(1), G(2), \cdots.$$

It is apparent that the series (4) may be written in the form

$$\sum_{k=1}^{n} \frac{1}{\lambda_k} x\left(1 - \frac{x^2}{1^2}\right)\left(1 - \frac{x^2}{2^2}\right)\cdots\left(1 - \frac{x^2}{(k-1)^2}\right), \tag{5}$$

where

$$\frac{1}{\lambda_k} = (-1)^{k-1}\overline{(k-1)!}^2\delta_s.$$

From the form (5) we see that $\sum_{k=1}^{n} \frac{1}{\lambda_k}$ is the coefficient of x in the polynomial of degree $2n - 1$ (at most) determined by the $2n + 1$ values of the given function $G(x)$ for $x = -n, \cdots, n$. The terms of the series $\sum_{k=1}^{n} \frac{1}{\lambda_k}$ have the following relation[3] to $G(x)$:

$$\sum_{j=1}^{n} \frac{\binom{2n}{n-j}}{\binom{2n}{n}}(-1)^{j-1}\frac{2G(j)}{j} = \sum_{k=1}^{n} \frac{1}{\lambda_k}. \tag{6}$$

We shall refer to (6) as *the finite de la Vallée-Poussin sum of the u-series* $\sum_{j=1}^{n}(-1)^{j-1}u_j$, where $u_j = \frac{2G(j)}{j}$. As a means of

[3] Rutledge, loc. cit., p. 262.

reference specifically to the right-hand side of (6) we shall speak of the λ-series.

A necessary and sufficient condition [4] for convergence of the Stirling series (4) is the convergence of (6), that is, the (VP) summability of $\sum(-1)^{i-1}u_i$.

However, we shall be concerned in this paper with the relation (6) for $n = 1, 2, 3, \cdots$ in divergent as well as convergent cases. *Our principal problem is the determination of λ-series by means of the (VP) summation process defined in* (6). When the λ-series converges the Stirling series (4) defines an integral function $I(x)$ such that $I(x) = G(x)$ for all integral values of x, including zero. It is, of course, not necessary that $I(x)$ and $G(x)$ be identical.

Sufficient conditions for convergence of the Stirling series determined by $G(x)$ to $G(x)$ have been given by Nörlund,[5] and others. Necessary conditions closely approximating these sufficient conditions have also been given by Nörlund.[6]

In the final section of this paper we shall state a necessary condition for (VP) summability of $\sum(-1)^{i-1}u_i$ which follows from Nörlund's necessary condition for the convergence of the Stirling series (4).

In another section (Sec. 6), we consider a group of u-series which are (VP) summable but not Cesàro summable of any order. The corresponding functions $G(x)$ are of the form $C \sinh \alpha x$ and have the interesting property of being integral-valued for integral x.

In the next section we shall obtain some general relations between the terms of corresponding u-series and λ-series.

2. Linear Relations between u_i and $\frac{1}{\lambda_k}$. We start with the relation, equivalent to (6),

[4] Rutledge, loc. cit., p. 263.

[5] Nörlund, Leçons sur les Series d'Interpolation, Paris, 1926. This monograph contains an extensive bibliography.

[6] Loc. cit.

$$\sum_{j=1}^{n} \frac{\binom{2n}{n-j}}{\binom{2n}{n}} (-1)^{j-1} u_j = \sum_{k=1}^{n} \frac{1}{\lambda_k}, \tag{7}$$

from which we have

$$\begin{aligned}
\frac{1}{\lambda_1} &= \frac{1}{1^2}\left[\frac{\binom{2}{0}}{\binom{2}{1}} 1^2 \cdot u_1\right], \\
\frac{1}{\lambda_2} &= \frac{1}{2^2}\left[\frac{\binom{4}{1}}{\binom{4}{2}} 1^2 \cdot u_1 - \frac{\binom{4}{0}}{\binom{4}{2}} 2^2 \cdot u_2\right], \\
\frac{1}{\lambda_3} &= \frac{1}{3^2}\left[\frac{\binom{6}{2}}{\binom{6}{3}} 1^2 \cdot u_1 - \frac{\binom{6}{1}}{\binom{6}{3}} 2^2 \cdot u_2 + \frac{\binom{6}{0}}{\binom{6}{3}} 3^2 \cdot u_3\right], \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned} \tag{8}$$

since (7) results by addition of the equations (8).

From the relation

$$\frac{G(j)}{j} = \frac{1}{\lambda_1} + \frac{1}{\lambda_2}\left(1 - \frac{j^2}{1^2}\right) + \frac{1}{\lambda_3}\left(1 - \frac{j^2}{1^2}\right)\left(1 - \frac{j^2}{2^2}\right) + \cdots \tag{9}$$

we have, by use of the identity

$$\begin{aligned}
&\left(1 - \frac{j^2}{1^2}\right)\left(1 - \frac{j^2}{2^2}\right)\cdots\left(1 - \frac{j^2}{(k-1)^2}\right) \\
&\qquad = (-1)^{k-1}\frac{k}{j}\binom{2k-1}{k-1}\binom{j+k-1}{2k-1},
\end{aligned} \tag{10}$$

the following equations which constitute the solution of the linear equations (8):

$$
\begin{aligned}
u_1 &= 2\left[\frac{1}{\lambda_1}\right], \\
-u_2 &= 2\left[-\frac{1}{\lambda_1}+\frac{2}{2}\binom{3}{1}\binom{3}{3}\frac{1}{\lambda_2}\right], \\
u_3 &= 2\left[\frac{1}{\lambda_1}-\frac{2}{3}\binom{3}{1}\binom{4}{3}\frac{1}{\lambda_2}+\frac{3}{3}\binom{5}{2}\binom{5}{5}\frac{1}{\lambda_3}\right], \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}
\tag{11}
$$

From (11) we have at once

$$
\begin{aligned}
G(j) &= \sum_{s=0}^{j-1}(-1)^s(s+1)\binom{2s+1}{s}\binom{j+s}{2s+1}\frac{1}{\lambda_{s+1}} \\
&= j\sum_{s=0}^{j-1}(-1)^s\binom{j-1}{s}\binom{j+s}{s}\frac{1}{\lambda_{s+1}},
\end{aligned}
\tag{12}
$$

since

$$
\binom{2s+1}{s}\binom{j+s}{2s+1} = \binom{j}{s+1}\binom{j+s}{s}.
$$

We shall have occasion to use the relation (12) also in the form

$$
u_j = \sum_{s=0}^{j-1}(-1)^s 2\binom{j-1}{s}\binom{j+s}{s}\frac{1}{\lambda_{s+1}}. \tag{13}
$$

3. Tabulated Results. Before proceeding with the detailed derivation of various λ-series, we present for convenience of reference a summary of results in tabular form.

In the following table various u-series and corresponding λ-series are presented in parallel columns. *It is to be understood that the relation* (6) *is valid for* $n = 1, 2, 3, \cdots$ *in each case.* Thus the first case tabulated implies the relations

$$
\begin{aligned}
&\frac{\binom{2}{0}}{\binom{2}{1}}2\cdot 1^2 = \frac{2}{4}\frac{2^2}{\binom{2}{1}}, \\
&\frac{\binom{4}{1}}{\binom{4}{2}}2\cdot 1^2 + \frac{\binom{4}{0}}{\binom{4}{2}}2\cdot 2^2 = \frac{2}{4}\frac{2^2}{\binom{2}{1}} + \frac{5}{8}\frac{2^4}{\binom{4}{2}}, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots
\end{aligned}
\tag{14}
$$

which we express briefly by writing

$$(VP)\sum_1^n 2j^2 = \sum_1^n \frac{3k-1}{4k}\frac{2^{2k}}{\binom{2k}{k}}. \tag{15}$$

TABLE OF FINITE (VP) SUMS

Oscillating $G(x)$

U-SERIES	λ-SERIES
Positive-term Divergent	*Positive-term Divergent*
$2[1^2+2^2+3^2+\cdots]$	$\sum \frac{3k-1}{4k}\frac{2^{2k}}{\binom{2k}{k}}$
$2[1+2+3+\cdots]$	$1+1+1+\cdots$
$2[1+1+1+\cdots]$	$\sum \frac{1}{2k}\frac{2^{2k}}{\binom{2k}{k}}$
$2\left[1+\frac{1}{2}+\frac{1}{3}+\cdots\right]$	$1+\frac{1}{2}+\frac{1}{3}+\cdots$
. . . . .	. . . . .
Positive-term Convergent	*Positive-term Convergent*
$2\left[1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots\right]$	$\sum \frac{1}{k^2}\left(\frac{2^{2k}}{\binom{2k}{k}}-1\right)$
$2\left[1+\frac{1}{2^3}+\frac{1}{3^3}+\cdots\right]$	$\sum \frac{1}{k^2}H(k),\quad H(k)=\sum_1^k \frac{1}{s}$
$2\left[1+\frac{1}{2^4}+\frac{1}{3^4}+\cdots\right]$	$\sum \frac{1}{k^2}J(k),$ $J(k)=\sum_1^k \frac{1}{s^2}\left(\frac{2^{2s}}{\binom{2s}{s}}-1\right)$

Non-Oscillating $G(x)$

U-SERIES	λ-SERIES
Alternating-term Convergent	*Positive-term Convergent*
$2\left[1-\frac{1}{2}+\frac{1}{3}-\cdots\right]$	$\sum\frac{1}{k(2k-1)}$
$2\left[1-\frac{1}{2^2}+\frac{1}{3^2}-\cdots\right]$	$1+\frac{1}{2^2}+\frac{1}{3^2}+\cdots$
$2\left[1-\frac{1}{2^3}+\frac{1}{3^3}-\cdots\right]$	$\sum\frac{1}{k^2}P(k),\quad P(k)=\sum_1^k\frac{1}{s(2s-1)}$
$2\left[1-\frac{1}{2^4}+\frac{1}{3^4}-\cdots\right]$	$\sum\frac{1}{k^2}R(k),\quad R(k)=\sum_1^k\frac{1}{s^2}$
. . . . .	. . . . .
Alternating-term Divergent	*Non-alternating Convergent*
$2[1-1+1-\cdots]$	$1+0+0+\cdots$
$2[1-2+3-\cdots]$	$1-\frac{1}{1\cdot3}-\frac{1}{3\cdot5}-\cdots$
$2[1-2^2+3^2-\cdots]$	$1-1+0+\cdots$
$2[1-2^3+3^3-\cdots]$	$1+\frac{7}{-1\cdot1\cdot3}+\frac{11}{1\cdot3\cdot5}+\frac{15}{3\cdot5\cdot7}+\cdots$
$2[1-2^4+3^4-\cdots]$	$1-5+4+0+\cdots$
. . . . .	. . . . .

Exponential[7] *$G(x)$*

U-SERIES	λ-SERIES
Alternating-term Divergent	*Alternating-term Convergent*
$\sum(-1)^{j-1}\frac{2\sinh j\alpha}{j}\cdot\frac{1}{2\sqrt{2}}$	$\sum(-1)^{k-1}\frac{2}{k}\frac{4^{k-1}}{\binom{2k}{k}}$

[7] For definition of constants α, β, γ, δ see page 150.

U-SERIES	λ-SERIES
Alternating-term Divergent	*Alternating-term Convergent*
$\sum (-1)^{j-1} \frac{2 \sinh j\beta}{j} \cdot \frac{2}{\sqrt{21}}$	$\sum (-1)^{k-1} \frac{2}{k} \frac{3^{k-1}}{\binom{2k}{k}}$
$\sum (-1)^{j-1} \frac{2 \sinh j\gamma}{j} \cdot \frac{1}{\sqrt{3}}$	$\sum (-1)^{k-1} \frac{2}{k} \frac{2^{k-1}}{\binom{2k}{k}}$
$\sum (-1)^{j-1} \frac{2 \sinh j\delta}{j} \cdot \frac{2}{\sqrt{5}}$	$\sum (-1)^{k-1} \frac{2}{k} \frac{1^{k-1}}{\binom{2k}{k}}$
$\sum (-1)^{j-1} \frac{2[\cosh j\alpha - 1]}{j^2} \cdot \frac{1}{2}$	$\sum (-1)^{k-1} \frac{2}{k^2} \frac{4^{k-1}}{\binom{2k}{k}}$
$\sum (-1)^{j-1} \frac{2[\cosh j\beta - 1]}{j^2} \cdot \frac{2}{3}$	$\sum (-1)^{k-1} \frac{2}{k^2} \frac{3^{k-1}}{\binom{2k}{k}}$
$\sum (-1)^{j-1} \frac{2[\cosh j\gamma - 1]}{j^2} \cdot 1$	$\sum (-1)^{k-1} \frac{2}{k^2} \frac{2^{k-1}}{\binom{2k}{k}}$
$\sum (-1)^{j-1} \frac{2[\cosh j\delta - 1]}{j^2} \cdot 2$	$\sum (-1)^{k-1} \frac{2}{k^2} \frac{1^{k-1}}{\binom{2k}{k}}$

We shall in the sections to follow derive the foregoing results, grouping cases the derivations of which are related.

4. Oscillating $G(x)$. Here $G(x)$ is called *oscillating* because $G(j)$ has the sign $(-1)^{j-1}$. In the cases to be considered in this section, the λ-series is convergent only when the u-series is convergent. We consider first divergent cases.

1. $u_j = \{(-1)^{j-1}2\}$. From the identity

$$\binom{2n}{n-1} + \binom{2n}{n-2} + \cdots + \binom{2n}{0} = \frac{1}{2}\left[2^{2n} - \binom{2n}{n}\right] \quad (16)$$

we have

$$\sum_{j=1}^{n} \frac{\binom{2n}{n-j}}{\binom{2n}{n}} (-1)^{j-1} \{(-1)^{j-1}2\} = \frac{2^{2n}}{\binom{2n}{n}} - 1. \quad (17)$$

From (17) we have, by reference to (6),

$$\frac{1}{\lambda_n} = \frac{2^{2n}}{\binom{2n}{n}} - \frac{2^{2n-2}}{\binom{2n-2}{n-1}} = \frac{1}{2n} \frac{2^{2n}}{\binom{2n}{n}}; \quad (18)$$

that is,

$$(VP) \sum_{1}^{n} 2 = \sum_{1}^{n} \frac{1}{2k} \frac{2^{2k}}{\binom{2k}{k}}. \quad (19)$$

By use of Stirling's approximation for $n!$ we have

$$\frac{2^{2k}}{\binom{2k}{k}} \leqq \sqrt{\pi}\sqrt{k}, \qquad \frac{1}{2k} \frac{2^{2k}}{\binom{2k}{k}} \leqq \frac{\sqrt{\pi}}{2} \frac{1}{\sqrt{k}}. \quad (20)$$

Thus the λ-series is divergent as must be the case when the u-series is properly divergent.

2. $u_j = \{(-1)^{j-1}2j^2\}$. From equations (8) we have the general relation (for any u_j)

$$\frac{1}{\lambda_k} = \frac{1}{k^2} \sum_{j=1}^{k} \frac{\binom{2k}{k-j}}{\binom{2k}{k}} (-1)^{j-1} j^2 u_j, \quad (21)$$

and with $u_j = \{(-1)^{j-1}2\}$, and $k = n$, we have by (19)

$$\sum_{j=1}^{n} \frac{\binom{2n}{n-j}}{\binom{2n}{n}} (-1)^{j-1} \{(-1)^{j-1}2j^2\} = \frac{n}{2} \frac{2^{2n}}{\binom{2n}{n}} \quad (22)$$

From (22) we now obtain, for the u-series 2,

$$\frac{1}{\lambda_n} = \frac{n}{2}\frac{2^{2n}}{\binom{2n}{n}} - \frac{n-1}{2}\frac{2^{2n-2}}{\binom{2n-2}{n-1}} = \frac{3n-1}{4n}\frac{2^{2n}}{\binom{2n}{n}}, \quad (23)$$

that is,

$$(VP)\sum_1^n 2j^2 = \sum_1^n \frac{3k-1}{4k}\frac{2^{2k}}{\binom{2k}{k}}. \quad (24)$$

The process here employed in obtaining (24) from (19) by use of (21) is applicable generally and will presently be stated in general form ((36) below).

3. $u_j = \left\{(-1)^{j-1}\frac{2}{j}\right\}$. Using equations (8) with the given value of u_j we obtain

$$\frac{1}{\lambda_k} = \frac{2}{k^2}\left[\frac{\binom{2k}{k-1}}{\binom{2k}{k}}1^2\frac{1}{1} + \frac{\binom{2k}{k-2}}{\binom{2k}{k}}2^2\frac{1}{2} + \cdots + \frac{\binom{2k}{0}}{\binom{2k}{k}}k^2\frac{1}{k}\right]$$

$$= \frac{2}{k^2}\frac{\frac{k}{2}\binom{2k}{k}}{\binom{2k}{k}} = \frac{1}{k}. \quad (25)$$

This follows from the identity [8]

$$\binom{2n}{n-1}1 + \binom{2n}{n-2}2 + \cdots + \binom{2n}{0}n = \frac{n}{2}\binom{2n}{n}. \quad (26)$$

Thus we have

$$(VP)\sum_1^n \frac{2}{j} = \sum_1^n \frac{1}{k}. \quad (27)$$

4. $u_j = \{(-1)^{j-1}2j\}$. Proceeding from (27) as in the deriva-

[8] Cairns, "Certain Properties of Binomial Coefficients," *Bulletin Amer. Math. Soc.*, vol. 2 (1919), p. 160.

tion of (24) from (19) we readily obtain ((36) below)

$$(VP)\sum_1^n 2j = \sum_1^n k^0 = 1 + 1 + 1 + \cdots. \tag{28}$$

It will be observed that the cases just discussed correspond (though not in order) to the first four cases tabulated, that is, to $(VP)\sum_1^n 2j^p$, $p = -1, 0, 1, 2$. It is evident that by repetition of the above procedure we may obtain corresponding results for any positive integral value of p. These results will fall in two groups, namely $p = 0, 2, \cdots$, $p = -1, 1, \cdots$, since $(VP)\sum j^2 u_j$ is obtainable from $(VP)\sum u_j$.

Let us now consider the convergent cases.

5. $u_j = \left\{(-1)^{j-1}\frac{2}{j^2}\right\}$. From equations (8)

$$\frac{1}{\lambda_k} = \frac{2}{k^2}\left[\frac{\binom{2k}{k-1}}{\binom{2k}{k}}\frac{1^2}{1^2} + \frac{\binom{2k}{k-2}}{\binom{2k}{k}}\frac{2^2}{2^2} + \cdots + \frac{\binom{2k}{0}}{\binom{2k}{k}}\frac{k^2}{k^2}\right]$$

$$= \frac{2}{k^2}\frac{\frac{1}{2}\left[2^{2k} - \binom{2k}{k}\right]}{\binom{2k}{k}} = \frac{1}{k^2}\left[\frac{2^{2k}}{\binom{2k}{k}} - 1\right], \tag{29}$$

by application of (16). Consequently

$$(VP)\sum_1^n \frac{2}{j^2} = \sum_1^n \frac{1}{k^2}\left(\frac{2^{2k}}{\binom{2k}{k}} - 1\right). \tag{30}$$

It may be noted that (19) may be derived from (30). Both depend fundamentally on the identity (16).

The limit function $G(x)$ in this case may be identified by virtue of its equivalence with a known cardinal series.[9] It is

[9] Whittaker, *Proc. Roy. Soc. Edin.*, vol. 35 (1915), p. 181. Concerning the general relation between the Stirling series for $G(x)$ and the cardinal series for $G(x)$ see Rutledge, loc. cit., p. 276.

$$G(x) = \frac{\sin \pi x}{\pi x}\left[\psi(x) - \psi(-x) + \frac{2}{x}\right]$$

$$= \frac{1}{x}\left[\frac{\sin \pi x}{\pi x} - \cos \pi x\right]. \tag{31}$$

6. $u_j = \left\{(-1)^{j-1}\frac{2}{j^2}\right\}$. From equations (8)

$$\frac{1}{\lambda_k} = \frac{1}{k^2}\sum_{j=1}^{k} \frac{\binom{2k}{k-j}}{\binom{2k}{k}}(-1)^{j-1}\left\{(-1)^{j-1}\frac{2}{j^2}\right\} j^2$$

$$= \frac{1}{k^2}\sum_{j=1}^{k} \frac{\binom{2k}{k-j}}{\binom{2k}{k}}\frac{2}{j}$$

$$= \frac{1}{k^2}\sum_{s=1}^{k}\frac{1}{s} = \frac{1}{k^2}H(k), \tag{32}$$

by application of (27).

Since the u-series is convergent and since (VP) summation does not alter its sum we have the interesting result

$$\sum_{1}^{\infty}\frac{2}{j^2} = \sum_{1}^{\infty}\frac{1}{k^2}\sum_{1}^{k}\frac{1}{s}. \tag{33}$$

The process by which (33) is obtained from (27) by use of equations (8) is applicable generally. This process and that described in connection with the u-series 2. and 4. above may be expressed in general form in the following manner.

Consider the following u- and λ-series, in which $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \frac{1}{\lambda_3}, \cdots$ are supposed known.

u-series	*λ-series*
$1^2 \cdot u_1 - 2^2 \cdot u_2 + 3^2 \cdot u_3 - \cdots$	$\frac{1}{\zeta_1} + \frac{1}{\zeta_2} + \frac{1}{\zeta_3} + \cdots$

$$u_1 - u_2 + u_3 - \cdots \qquad \frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3} + \cdots \quad (34)$$

$$\frac{u_1}{1^2} - \frac{u_2}{2^2} + \frac{u_3}{3^2} - \cdots \qquad \frac{1}{\xi_1} + \frac{1}{\xi_2} + \frac{1}{\xi_3} + \cdots$$

From equations (8), and from the definition (6) we then have the following general relations:

$$\frac{1}{\xi_k} = \frac{1}{k^2} \sum_{s=1}^{k} \frac{1}{\lambda_s}, \tag{35}$$

$$\frac{1}{\zeta_k} = k^2 \frac{1}{\lambda_k} - (k-1)^2 \frac{1}{\lambda_{k-1}}. \tag{36}$$

It is apparent that theoretically formulas (35) and (36) enable us to obtain the λ-series for the u-series

$$\sum 2j^p \text{ for } p = \cdots, -3, -2, -1, 0, 1, 2, 3, \cdots,$$

given any two successive ones of these as a basis, but extension to higher negative powers is subject to the disadvantage involved in our inability to express $\sum_{s=1}^{k} \frac{1}{s}$ more simply. We encounter precisely this difficulty again in the next u-series to be considered.

7. $u_j = \left\{ (-1)^{j-1} \frac{2}{j^4} \right\}$. By application of (35) to (30) we have at once

$$\frac{1}{\lambda_k} = \frac{1}{k^2} \sum_{s=1}^{k} \frac{1}{s^2} \left(\frac{2^{2s}}{\binom{2s}{s}} - 1 \right) = \frac{1}{k^2} J(k). \tag{37}$$

5. Non-oscillating $G(x)$. In the cases to be considered in this section, $G(j)$ is positive for positive j, and equal to an odd or even positive or negative power of j.

The λ-series is convergent in all cases. In the positive odd power cases $G(j)$ reduces to x, x^3, x^5, $\cdots$. The positive even power cases, that is the cases where $G(j) = j^0$, $G(j) = j^2$, $G(j) = j^4$, $\cdots$, are particularly interesting in view of the fact that the limit function $G(x)$ is an *odd* integral function.

We consider first the positive odd power cases, $G(j) = j$, $G(j) = j^3$, $\cdots$.

1. $u_j = 2, 2j^2, 2j^4, \cdots$. Applying definition (6) with $u_j = 2$ we have

$$\frac{\binom{2n}{n-1}}{\binom{2n}{n}} 2 - \frac{\binom{2n}{n-2}}{\binom{2n}{n}} 2 + \frac{\binom{2n}{n-3}}{\binom{2n}{n}} - \cdots + (-1)^{n-1} \frac{\binom{2n}{0}}{\binom{2n}{n}} 2 = 1. \quad (38)$$

This follows from the identity

$$\binom{2n}{n-1} - \binom{2n}{n-2} + \binom{2n}{n-3} + \cdots (-1)^{n-1} \binom{2n}{0} = \frac{1}{2}\binom{2n}{n}. \quad (39)$$

From (38) and (6) we see that

$$\frac{1}{\lambda_k} = 1, \quad k = 1; \quad \frac{1}{\lambda_k} = 0, \quad k > 1; \quad (40)$$

and hence $G(x) = x$.

By successive application of (36), with (40) as a basis, and $u_j = 2j^{2p}$, we find λ-series with corresponding $G(x)$ for $p = 0, 1, 2, \cdots$, as follows:

p	λ-*series*	$G(x)$	
0	$1 - 0 + 0 + \cdots$	x	
1	$1 - 1 + 0 + \cdots$	x^3	
2	$1 - 5 + 4 + 0 + \cdots$	x^5	
3	$1 - 21 + 56 - 36 + 0 + \cdots$	x^7	(41)
.	.	.	
.	.	.	
.	.	.	

Thus the function $G(x)$ here reduces as stated above to a rational function, namely an odd power of x.

It is apparent that the process used for determining the λ-series above is a recursion formula by means of which we may rapidly compute the central differences of the functions x, x^3, x^5, x^7, $\cdots$.

We may also note that by direct application of (6) these relations give rise to the following identity:

$$\binom{2n}{n-1}1^{2p} - \binom{2n}{n-2}2^{2p} + \binom{2n}{n-3}3^{2p} - \cdots = 0, \quad \text{for } n > p. \tag{42}$$

Next we consider the positive even power cases, $G(j) = j^0$, $G(j) = j^2$, $\cdots$.

2. $u_j = \frac{2}{j}$, $2j$, $2j^2$, $\cdots$. From equations (8) with $u_j = \frac{2}{j}$ we have

$$\frac{1}{\lambda_k} = \frac{2}{k^2}\frac{\binom{2k}{k-1}}{\binom{2k}{k}}1^2\frac{1}{1} - \frac{\binom{2k}{k-2}}{\binom{2k}{k}}2^2\cdot\frac{1}{2} + \cdots + (-1)^{k-1}\frac{\binom{2k}{0}}{\binom{2k}{k}}k^2\frac{1}{k}$$

$$= \frac{2}{k^2}\frac{\binom{2k-2}{k-1}}{\binom{2k}{k}} = \frac{2}{k^2}\cdot\frac{k}{2}\frac{1}{2k-1} = \frac{1}{k(2k-1)}. \tag{43}$$

This follows from the identity

$$\binom{2n}{n-1}1 - \binom{2n}{n-2}2 + \cdots + (-1)^{n-1}\binom{2n}{0}n = \binom{2n-2}{n-1}, \tag{44}$$

which is derived by comparing coefficients of x^{n-1} in

$$\frac{1}{(1+x)^2}\cdot(1+x)^{2n} = (1-2x+3x^2-\cdots) \times\left[\binom{2n}{0}+\binom{2n}{1}x+\binom{2n}{2}x^2+\cdots\right].$$

By successive application of (36) with (43) as a basis, we easily obtain

$$(VP)\sum_1^n(-1)^{i-1}2j = \sum_1^n \frac{-1}{(2k-3)(2k-1)} \tag{45}$$

$$(VP)\sum_1^n(-1)^{i-1}2j^3 = \sum_1^n \frac{4k-1}{(2k-5)(2k-3)(2k-1)} \tag{46}$$

. . . . . .

In the first mentioned of these even power cases (43) the limit function $G(x)$ has been previously identified by Rutledge [10] as follows:

$$G(x) = \frac{\sin \pi x}{\pi x}\left[-\psi(x)-\psi(-x)+\psi\left(\frac{x}{2}\right) +\psi\left(-\frac{x}{2}\right)+2\log 2\right]x.$$

This function takes the value $+1$ for every positive integral value of x and the value -1 for every negative integral value of x. It is defined by its expansion in the Stirling series (4) or (5). The λ-series (43) is convergent to the value $2\log 2$, which is, therefore, the coefficient of x in the Maclaurin expansion of $G(x)$.

We now turn to the consideration of $G(j)=\frac{1}{j}$, $G(j)=\frac{1}{j^3}$, $\cdots$.

3. $u_j=\frac{2}{j^2}, \frac{2}{j^4}, \frac{2}{j^6}, \cdots$. From equations (8) for $u_j=\frac{2}{j^2}$ we have, by (39) or by application of (35) to (38),

$$\frac{1}{\lambda_k}=\frac{1}{k^2}. \tag{47}$$

[10] Rutledge, loc. cit., p. 271.

This function also has been identified by Rutledge [11] as follows:

$$G(x) = \frac{\sin \pi x}{\pi x}\left[\psi(x) - \psi(-x) - \psi\left(\frac{x}{2}\right) + \psi\left(-\frac{x}{2}\right) - \frac{2}{x}\right]$$
$$= \frac{1}{x}\left[1 - \frac{\sin \pi x}{\pi x}\right]. \tag{48}$$

The λ-series is convergent to $\frac{\pi^2}{3!}$.

To determine $\frac{1}{\lambda_k}$ for $u_j = \frac{2}{j^4}$ it is necessary to use (35) with (47). We then have

$$\frac{1}{\lambda_k} = \frac{1}{k^2}\sum_{s=1}^{k}\frac{1}{s^2} = \frac{1}{k^2}R(k). \tag{49}$$

The (VP) sum $(VP)\sum(-1)^{j-1}u_j$ may in this case be conveniently expressed by the use of symmetric functions of the squared reciprocals of the first n integers. These symmetric functions have well-known limits as n becomes infinite, which may be readily derived from the following relations:

It is obvious and well-known that the product

$$\frac{\sinh \pi x}{\pi x} = \left(1 + \frac{x^2}{1^2}\right)\left(1 + \frac{x^2}{2^2}\right)\left(1 + \frac{x^2}{3^2}\right)\cdots \tag{50}$$

may be regarded as a convergent double series of positive terms,

$$1 + \sum_1^\infty \frac{1}{r_1^2}x^2 + \sum_1^\infty \frac{1}{r_1^2 r_2^2}x^4 + \sum_1^\infty \frac{1}{r_1^2 r_2^2 r_3^2}x^6 \cdots, \tag{51}$$

from which by comparison with

$$\frac{\sinh \pi x}{\pi x} = 1 + \frac{\pi^2 x^2}{3!} + \frac{\pi^4 x^4}{5!} + \frac{\pi^6 x^6}{7!} + \cdots. \tag{52}$$

It follows that

$$\sum_1^\infty \frac{1}{r^2} = \frac{\pi^2}{3!}, \qquad \sum_1^\infty \frac{1}{r_1^2 r_2^2} = \frac{\pi^4}{5!}, \qquad \sum_1^\infty \frac{1}{r_1^2 r_2^2 r_3^2} = \frac{\pi^6}{7!}\cdots. \tag{53}$$

[11] Rutledge, loc. cit., p. 274.

Furthermore,[12] both for n finite and for n infinite,

$$\left(\Sigma \frac{1}{r_1^2}\right)\left(\Sigma \frac{1}{r_1^2}\right) = 2\Sigma \frac{1}{r_1^2 r_2^2} + \Sigma \frac{1}{r_1^4},$$

$$\left(\Sigma \frac{1}{r_1^2 r_2^2}\right)\left(\Sigma \frac{1}{r_1^2}\right) = 3\Sigma \frac{1}{r_1^2 r_2^2 r_3^2} + \Sigma \frac{1}{r_1^4 r_2^2}, \quad (54)$$

. . . . . .

Now from (49) we have

$$(VP)\sum_1^n (-1)^{i-1}\frac{2}{j^4} = \sum_1^n \frac{1}{k^2} R(k)$$

$$= \left(\frac{1}{1^2}\right)\frac{1}{1^2} + \left(\frac{1}{1^2}+\frac{1}{2^2}\right)\frac{1}{2^2} + \left(\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}\right)\frac{1}{3^2} + \cdots$$

$$= \sum_1^n \frac{1}{r_1^2 r_2^2} + \sum_1^n \frac{1}{r_1^4}, \quad (55)$$

and, from (53) and (54)

$$\sum_1^\infty \frac{1}{\lambda_k} = \frac{\pi^4}{5!} + \frac{\pi^4}{90} = \frac{7\pi^4}{360}. \quad (56)$$

Similarly, for $u_j = \frac{2}{j^6}$ we find

$$\sum_1^n \frac{1}{\lambda_k} = \sum_1^n \frac{1}{r_1^2 r_2^4} + \sum_1^n \frac{1}{r_1^2 r_2^2 r_3^2} + \sum_1^n \frac{1}{r_1^6} \quad (57)$$

and hence from (53) and (54) and with

$$\sum_1^\infty \frac{1}{r_1^6} = \frac{\pi^6}{945},$$

we obtain the following:

$$\sum_1^\infty \frac{1}{\lambda_k} = \frac{31\pi^6}{15120}. \quad (58)$$

[12] Rutledge, "Convergent Interpolation Coefficients with Convergent Sum," *Jour. of Math. and Phys.*, vol. 6 (1927), p. 248.

The results (56) and (58) are in agreement with [13]

$$\sum_{1}^{\infty}(-1)^{r-1}\frac{2}{r^4}=\frac{7\pi^4}{360},\qquad \sum_{1}^{\infty}(-1)^{r-1}\frac{2}{r^6}=\frac{31\pi^6}{15120}.$$

Finally, we consider negative even power cases, $G(j)=\frac{1}{j^2}$ $G(j)=\frac{1}{j^4}$, $\cdots$.

4. $u_j=\frac{2}{j^2},\frac{2}{j^3},\cdots$. By successive application of (35) with (43) as a basis, we easily obtain

$$(VP)\sum(-1)^{j-1}\frac{2}{j^2}=\sum_{k=1}^{n}\frac{1}{k^2}\sum_{s=1}^{k}\frac{1}{s(2s-1)}=\sum_{1}^{n}\frac{1}{k^2}P(k), \tag{59}$$

$$(VP)\sum_{1}^{n}(-1)^{j-1}\frac{2}{j^3}=\sum_{k=1}^{n}\frac{1}{k^2}\sum_{s=1}^{k}\frac{1}{s^2}P(s)=\sum_{1}^{n}\frac{1}{k^2}Q(k). \tag{60}$$

. . . . . .

6. Exponential u-series. The Stirling series expression of the function e^{hz} converges to e^{hz} for all positive values of $h\leqq 2\sinh^{-1}1$. This was shown by Runge [14] who established (for the convergence to e^{hz} of the Gauss expansion) the condition $h<2\sinh^{-1}1$. A slight extension of his argument shows that if $h=2\sinh^{-1}1$ the Stirling series still converges to e^{hz}. For greater values of h this series diverges.

We shall now consider this and related Stirling series, obtaining in this manner series of the form $\sum(-1)^{j-1}u_j$ for which the terms u_j are positive and increasing exponentially as j increases. These series are consequently not Cesàro summable of any order, but do have convergent λ-series, that is, are de la Vallée-Poussin summable.

The mean central differences required for the Stirling expansion (4) of e^{hz} may be placed in the following form:

[13] Knopp, *Theory and Application of Infinite Series*, pp. 237–239.

[14] *Zeitschrift für Mathematik und Physik*, 46 (1901), pp. 224–243.

Mean Central Differences of e^{hz}

	$h=\alpha$	$h=\beta$	$h=\gamma$	$h=\delta$
1	1	1	1	1
UV	$\frac{\sqrt{2}}{2}\cdot 4$	$\frac{\sqrt{21}}{6}\cdot 3$	$\frac{\sqrt{3}}{2}\cdot 2$	$\frac{\sqrt{5}}{2}\cdot 1$
V	4	3	2	1
UV^2	$\frac{\sqrt{2}}{2}\cdot 4^2$	$\frac{\sqrt{21}}{5}\cdot 3^2$	$\frac{\sqrt{3}}{2}\cdot 2^2$	$\frac{\sqrt{5}}{2}\cdot 1^2$
V^2	4^2	3^2	2^2	1^2
UV^3	$\frac{\sqrt{2}}{2}\cdot 4^3$	$\frac{\sqrt{21}}{6}\cdot 3^3$	$\frac{\sqrt{3}}{2}\cdot 2^3$	$\frac{\sqrt{5}}{2}\cdot 1^3$
V^3	4^3	3^3	2^3	1^3
⋮	⋮	⋮	⋮	⋮

$U=\frac{1}{2}\coth\frac{h}{2}$, $V=4\sinh^2\frac{h}{2}$; $\cosh\alpha=3$, $\cosh\beta=\frac{5}{2}$, $\cosh\gamma=2$, $\cosh\delta=\frac{3}{2}$; $\sinh\alpha=2\sqrt{2}$, $\sinh\beta=\frac{\sqrt{21}}{2}$, $\sinh\gamma=\sqrt{3}$, $\sinh\delta=\frac{\sqrt{5}}{2}$.

We then have

$$e^{\alpha x}=1+\sinh\alpha x+(\cosh\alpha x-1)$$
$$=1+$$
$$\frac{\sqrt{2}}{2}\left[\frac{4}{1!}x-\frac{4^2 1^2}{3!}x\left(1-\frac{x^2}{1^2}\right)+\frac{4^2 1^2 2^2}{5!}x\left(1-\frac{x^2}{1^2}\right)\left(1-\frac{x^2}{2^2}\right)+\cdots\right]+\frac{4}{2!}x^2-\frac{4^2 1^2}{4!}x^2\left(1-\frac{x^2}{1^2}\right)+\frac{4^2 1^2 2^2}{6!}x^2\left(1-\frac{x^2}{1^2}\right)\left(1-\frac{x^2}{2^2}\right)+\cdots. \quad (61)$$

By reference to (3), (5) and (6) we now have from (61)

$$(VP)\sum_1^n (-1)^{j-1}\frac{2\sinh j\alpha}{j} = \sum_1^n (-1)^{k-1}\frac{1}{k}\frac{4^k}{\binom{2k}{k}}\cdot\sqrt{2}, \quad (62)$$

or dividing each side by $\sinh\alpha = 2\sqrt{2}$,

$$(VP)\sum_1^n (-1)^{j-1}\frac{2\sinh j\alpha}{j}\cdot\frac{1}{2\sqrt{2}} = \sum_1^n (-1)^{k-1}\frac{2}{k}\frac{4^{k-1}}{\binom{2k}{k}}. \quad (63)$$

An obvious modification of (61) yields the similar results for $\sinh j\beta$, $\sinh j\gamma$, $\sinh j\delta$ as tabulated in Sec. 3.

Since $\alpha = 2\sinh^{-1}1 = \log(1+\sqrt{2})^2$, we see that

$$\frac{1}{2\sqrt{2}}\sinh j\alpha = \sum_{s=0}^{j-1}\frac{1}{2}\binom{2j}{2s+1}2^s, \quad (64)$$

defining a sequence of integers. Similar expressions for the other three functions are readily obtained. The recursion formula

$$C[\sinh jh + \sinh(j-2)h] = (2\cosh h)\,C\sinh(j-1)h, \quad (65)$$

in the form

$$C\sinh jh = (2\cosh h)\,C\sinh(j-1)h - C\sinh(j-2)h, \quad (66)$$

when $2\cosh h = 6, 5, 4, 3$ for $h = \alpha, \beta, \gamma, \delta$, respectively, permits ready computation of the following sequences of integral values, and at the same time shows that no other value of $h < \alpha$ can yield a similar integral-valued hyperbolic sine function.

$$\begin{aligned}
\frac{1}{2\sqrt{2}}\sinh j\alpha &= 1, 6, 35, 204, 1189, 6930, \cdots,\\
\frac{2}{\sqrt{21}}\sinh j\beta &= 1, 5, 24, 115, 551, 2640, \cdots,\\
\frac{1}{\sqrt{3}}\sinh j\gamma &= 1, 4, 15, 56, 209, 780, \cdots,\\
\frac{2}{\sqrt{5}}\sinh j\delta &= 1, 3, 8, 21, 55, 144, \cdots.
\end{aligned} \quad (67)$$

In like manner we obtain from

$$C\cosh jh = (2\cosh h)\,C\cosh(j-1)h - C\cosh(j-2)h, \tag{68}$$

$$\begin{aligned}
&\tfrac{1}{2}(\cosh j\alpha - 1) = 1, 8, 49, 288, 1681, 9800, \cdots,\\
&\tfrac{2}{3}(\cosh j\beta - 1) = 1, 7, 36, 175, 841, 4032, \cdots,\\
&1(\cosh j\gamma - 1) = 1, 6, 25, 96, 361, 1350, \cdots,\\
&2(\cosh j\delta - 1) = 1, 5, 16, 45, 121, 320, \cdots.
\end{aligned} \tag{69}$$

The summation

$$(VP)\sum_1^n (-1)^{j-1}\frac{2[\cosh j\alpha - 1]}{j^3}\cdot\frac{1}{2} = \sum_1^n (-1)^{k-1}\frac{2}{k^2}\frac{4^{k-1}}{\binom{2k}{k}} \tag{70}$$

and the similar results for $\cosh j\beta$, $\cosh j\gamma$, $\cosh j\delta$ as tabulated in Sec. 3 are also an immediate consequence of the Stirling series (61) and its modifications.

By use of $\alpha = \log(1+\sqrt{2})^2$, as in the derivation of (64), we have

$$\tfrac{1}{2}(\cosh j\alpha - 1) = \sum_{s=0}^{j-1}\binom{2j}{2s+2}2^s \tag{71}$$

as an explicit definition of the first sequence of integers in (69).

It is evident that a generalization of the foregoing, with results analogous to (63) and (70) in which 4^{k-1} in the λ-series is replaced by $5^{k-1}, 6^{k-1}, 7^{k-1}, \cdots$, is obtained by taking $h = \cosh^{-1}\frac{7}{2}$, $\cosh^{-1}4$, $\cosh^{-1}\frac{9}{2}, \cdots$, respectively. However, since (compare (20))

$$\frac{5^k}{\binom{2k}{k}} \leqq \sqrt{\pi}\sqrt{k}\left(\frac{5}{4}\right)^k, \tag{72}$$

with similar asymptotic relations for all values of $h > \alpha$, we see that for $h > \alpha$ the finite (VP) sums

$$(VP)\sum_1^n (-1)^{j-1}\frac{2\sinh jh}{j}, \tag{73}$$

$$(VP)\sum_{1}^{n}(-1)^{j-1}\frac{2[\cosh jh-1]}{j}, \tag{74}$$

fail to have limits as n becomes infinite.

7. Necessary Conditions for (VP) Summability. The (VP) sum (63), that is,

$$(VP)\sum_{1}^{n}(-1)^{j-1}\frac{2}{j}\left\{\frac{1}{2\sqrt{2}}\sinh j\alpha\right\}=\sum_{1}^{n}(-1)^{k-1}\frac{2}{k}\frac{4^{k-1}}{\binom{2k}{k}} \tag{75}$$

has $(-1)^{k-1}\frac{\sqrt{\pi}}{2}\frac{1}{\sqrt{k}}$ as the asymptotic equivalent of the terms of the λ-series on the right. This λ-series is thus only conditionally convergent and suggests the divergent λ-series

$$1-1+1-1+1-\cdots, \tag{76}$$

which we shall now investigate.

Using equation (12) with $\frac{1}{\lambda_{s+1}}=(-1)^s$ and denoting $G(j)$ in this case by $L(j)$ we have

$$L(j)=j\sum_{s=0}^{j-1}\binom{j-1}{s}\binom{j+s}{s}. \tag{77}$$

By equating the coefficients of x^{j-1} in the identity

$$(1-x)^{-j-1}\left(1+\frac{1}{x}\right)^{j-1}x^{j-1}=\left(\frac{1+x}{1-x}\right)^{j-1}\left(\frac{1}{1-x}\right)^{2}$$

we obtain an equivalent form of $L(j)$, namely

$$L(j)=j\sum_{s=0}^{j-1}\binom{j-1}{s}\binom{j}{s+1}2^s. \tag{78}$$

For $j=1, 2, \cdots$, we have from (77) or (78)

$$L(j)=1, 8, 57, 384, 2505, 16008, 100849, \cdots. \tag{79}$$

From (64) and (78) we have

$$\frac{L(j)}{j\left(\frac{1}{2\sqrt{2}}\sinh j\alpha\right)} = \frac{\sum_{s=0}^{j-1}\binom{j-1}{s}\binom{j}{s+1}2^s}{\sum_{s=0}^{j-1}\frac{1}{2}\binom{2j}{2s+1}2^s}. \tag{80}$$

The ratio of the coefficients of 2^s in numerator and denominator on the right is

$$\frac{\binom{j}{s+1}\binom{j-1}{s}}{\frac{1}{2}\binom{2j}{2s+1}} = \frac{\binom{2j-2s}{j-s}}{\binom{2j}{j}}\cdot\frac{j-s}{j}. \tag{81}$$

For $s = 0$, the ratio is

$$\frac{\binom{2j}{j}\binom{1}{1}}{\binom{2j}{j}}\cdot\frac{j}{j} = 1. \tag{82}$$

Let the ratio for arbitrary s be written in the form

$$1 \cdot r_1 \cdot r_2 \cdot r_3 \cdots r_s.$$

Then

$$r_s = \frac{(j-s+1)^2}{[2(j-s+1)][2(j-s+1)-1]}\cdot\frac{4s+2}{s+1}\cdot\frac{j-s}{j-s+1}, \qquad s = 1, 2, \cdots, j-1. \tag{83}$$

Placing $j - s + 1 = n$, $0 < \epsilon = \frac{2}{s+1} \leq 1$, we have

$$r_s = \frac{(4-\epsilon)n^2}{4n^2-2n}\cdot\frac{n-1}{n} = \frac{(4-\epsilon)(n-1)}{4n-2}$$

$$= \frac{4n-2-(2+\epsilon n-\epsilon)}{4n-2}. \tag{84}$$

But $n \geq 2$, $\epsilon > 0$, and therefore

$$2 + \epsilon(n-1) > 0 \tag{85}$$

and hence $r_s < 1$.

Hence the ratio of the coefficients of 2^s in the numerator and denominator respectively of (80) which is equal to 1 for $s = 0$ diminishes monotonically with increasing s to $1/j$. The ratio of coefficients of 2^s in

$$\frac{L(j)}{\frac{1}{2\sqrt{2}}\sinh j\alpha} = \frac{j\sum_{s=0}^{j-1}\binom{j}{s+1}\binom{j-1}{s}2^s}{\sum_{s=0}^{j-1}\frac{1}{2}\binom{2j}{2s+1}2^s} \cdots, \tag{86}$$

then diminishes monotonically from j to 1. Therefore, from (80) and (86) we have the relation, for $j > 1$,

$$\frac{1}{2\sqrt{2}}\sinh j\alpha < L(j) < j\frac{1}{2\sqrt{2}}\sinh j\alpha. \tag{87}$$

The inequality (87) expresses a relation between *integers*, namely

$$\begin{aligned} j\frac{1}{2\sqrt{2}}\sinh j\alpha &= 1, 12, 105, 816, 5945, \cdots, \\ L(j) &= 1;\ 8,\ 57, 384, 2505. \cdots, \\ \frac{1}{2\sqrt{2}}\sinh j\alpha &= 1,\ 6,\ 35, 204, 1189, \cdots, \end{aligned} \tag{88}$$

and hence the addition to 6, 35, $\cdots$, in the third of the sequences (88), of the values

$$\frac{1}{2\sqrt{2}}\frac{e^{-2\alpha}}{2} = .005^+, \qquad \frac{1}{2\sqrt{2}}\frac{e^{-3\alpha}}{2} = .0008^+, \cdots, \tag{89}$$

respectively, does not disturb this inequality.

Therefore we have, for $j > 1$,

$$\frac{1}{4\sqrt{2}}e^{j\alpha} < L(j) < \frac{1}{4\sqrt{2}}je^{j\alpha}. \tag{90}$$

Now a necessary condition for convergence of Stirling's series for $G(x)$, where $G(x)$ is an odd function and where $G(x) > 0$,

for $x > 0$, is that $G(j)$ shall satisfy an inequality of the form

$$G(j) < C\sqrt{j}\, e^{j\alpha}, \tag{91}$$

where α has the same value as in (87) and (90). This condition is contained in a set of necessary conditions on $G(x)$ as stated by Nörlund.[15] It may here be interpreted as a necessary condition for (VP) summability of $\sum (-1)^{j-1} \frac{2G(j)}{j}$.

The precise relation between $L(j)$ and the function $\sqrt{j}\, e^{j\alpha}$ remains a subject for further investigation.

[15] Nörlund, loc. cit., p. 33.

GEODESICS OF PFAFFIANS[1]

By P. Franklin and C. L. E. Moore

1. *Introduction.* For the greater part of this paper, our basic geometric configuration is a single Pfaffian in Euclidean 3-space. Thus at each point of space the differential vectors tangent to the configuration lie in a plane, and arc lengths are measured in the usual way. Analytically,[2] such a Pfaffian is defined by a linear differential form

$$Adx + Bdy + Cdz = 0, \tag{1.1}$$

with coefficients functions of x, y, z and the Euclidean space in which it is embedded is characterized by the quadratic form

$$dx^2 + dy^2 + dz^2 = ds^2. \tag{1.2}$$

We are at first concerned with the curves in the Pfaffian of shortest length. That is, given two points in space which may be joined by curves lying in the Pfaffian, we seek the curve in the Pfaffian joining the points for which the length of arc between the points is a minimum. We set up the differential equations for these curves, integrate them for certain simple special cases, and study some of their local geometric properties for the general case.

In classical surface theory, the geodesic lines may either be defined as lines of shortest length, or as self parallel curves. The lines of shortest length just described are one analogue of geodesics for the Pfaffian. Another set of curves of the Pfaffian analogous to geodesics are the self parallel curves, defined as curves whose principal normal at each point coincides with the

[1] Presented to the American Mathematical Society.

[2] For our previous studies of such Pfaffians, cf. (I) The Geometry of Algebraic Pfaffians, vol. X, p. 19, and (II) Pfaffians in Parametric Form, vol. X, p. 95, this Journal.

PHYSICS LIBRARY

normal to the Pfaffian at that point. Such curves have been recently defined, for a more general configuration, by Schouten,[3] and were previously obtained by Dall' Acqua [4] from a different point of view as the solution of a minimization problem. While there are some points of similarity between the differential equations of the lines of shortest length, and the self parallel curves, they must, in general, be distinguished. For the general Pfaffian, there are a double infinity of lines of shortest length through each point, but only a single infinity of self parallel curves. Moreover, for non-integrable Pfaffians, the latter are not a sub-set of the former. In fact, we prove that a curve with non-zero curvature is simultaneously a line of shortest length and a self parallel curve when and only when the Pfaffian is integrable at each of its points. Both types of curve are of interest from the point of view of applications to mechanics. The lines of shortest length are the positions of equilibrium of stretched strings constrained to lie in the Pfaffian, while the self parallel curves are the trajectories for a suitably selected non-holonomic mechanical system.[5]

After studying these two types of curves for the single Pfaffian in Euclidean 3-space, we indicate briefly the corresponding types of curves for the configuration formed of m Pfaffians in an n-dimensional Riemannian space.

2. *Lines of Shortest Length, First Derivation.* The lines of shortest length of the Pfaffian:

$$Adx + Bdy + Cdz = 0 \tag{2.1}$$

are curves such that at each point this relation is satisfied, and such that for any segment of the curve the integral

$$\int ds \tag{2.2}$$

is a minimum. More precisely, we mean a relative minimum in

[3] Schouten, *Mathematische Zeitschrift*, vol. 30, 1929, p. 168; cf. also Vranceanu, *Paris Comptes Rendus*, vol. 183 (1926), p. 825 and p. 1083.

[4] Dall' Acqua, *Annali di Matematica*, vol. VI, Series II (1901), p. 11.

[5] First observed by A. Voss, *Math. Annalen*, vol. 25 (1885), p. 283.

the small, so that the points must be sufficiently near together, and the varied curves may have to be restricted to some neighborhood of the minimizing curve. Naturally, our curve is only compared with other curves lying in the Pfaffian.

We obtain the analytic conditions for the minimum problem by a method due to Lagrange.[6] That is, we introduce a multiplier $L(t)$, t being a parameter specifying points of the curve, and minimize:

$$I = \int [ds + L(t)(Adx + Bdy + Cdz)]. \tag{2.3}$$

This may be also written:

$$I = \int [\sqrt{x'^2 + y'^2 + z'^2} + L[Ax' + By' + Cz')]dt. \tag{2.31}$$

We shall obtain the differential equations for the lines of shortest length by equating the first variation of this integral, I, to zero. The curves satisfying these equations could be shown to be actual minimizing curves for sufficiently small arcs by a procedure which we briefly outline.[7] First, partition space into cells by a cubical network. Second, approximate the given Pfaffian by one with coefficients constant in each cell of the network, and approximate curved arcs in the original Pfaffian by broken lines made up of straight segments in each cell of the network. Since these polygonal lines depend on only a finite number of parameters, there exists a polygonal line of minimum length joining two given points for each size of network. Finally, prove that when the cells of the network close down these minimizing polygonal lines approach curves in the Pfaffian which are lines of shortest length. In place of supplying the details of this procedure, we shall content ourselves here with the intuitive argument that our curves are stationary, and hence maximum, minimum, or mini-max curves. But, in any neighborhood of a curve in the Pfaffian, an approximating zigzag curve can be

[6] Cf., e.g., Bliss, *American Journal of Math.*, vol. 52 (1930), p. 674.

[7] A study from the calculus of variations point of view is announced by I. Schoenberg, *Bull. A. M. S.*, vol. 37, 1931, p. 345.

drawn which has a greater length, so that the possibility of a maximum or of a mini-max is excluded.

For (2.31) the first variation is found to be:

$$\delta I = \int (P\delta x + Q\delta y + R\delta z)dt, \tag{2.32}$$

when we keep the end points fixed, where

$$\begin{aligned} P &= d/dt(x'/s' + AL) - L(A_x x' + B_x y' + C_x z'), \\ Q &= d/dt(y'/s' + BL) - L(A_y x' + B_y y' + C_y z'), \\ R &= d/dt(z'/s' + CL) - L(A_z x' + B_z y' + C_z z'). \end{aligned} \tag{2.33}$$

Since the problem places no restrictions on the parameter t, we may simplify these expressions by taking t equal to s. We continue to use primes to denote differentiation with respect to the parameter, so that from now on they denote s differentiation. The vanishing of the first variation of I is equivalent to the vanishing of P, Q, R, so that our conditions become:

$$\begin{aligned} P &= x'' + AL' + Ly'(A_y - B_x) - Lz'(C_x - A_z) = 0, \\ Q &= y'' + BL' + Lz'(B_z - C_y) - Lx'(A_y - B_x) = 0, \\ R &= z'' + CL' + Lx'(C_x - A_z) - Ly'(B_z - C_y) = 0. \end{aligned} \tag{2.4}$$

In deriving the equations (2.4) we used equations (2.3) in which the multiplier of ds was unity.[8] If this multiplier vanished, we would have in place of the set (2.41), similar equations without the second derivatives appearing, namely:

$$\begin{aligned} AL' + Ly'(A_y - B_x) - Lz'(C_x - A_z) &= 0, \\ BL' + Lz'(B_z - C_y) - Lx'(A_y - B_x) &= 0, \\ CL' + Lx'(C_x - A_z) - Ly'(B_z - C_y) &= 0. \end{aligned} \tag{2.41}$$

We will show that, when we take account of the conditions

$$Ax' + By' + Cz' = 0, \tag{2.42}$$

and

$$x'^2 + y'^2 + z'^2 = 1, \tag{2.43}$$

[8] See (6) l. c., p. 687–688. Equations (2.4) apply to what Bliss calls the "normal" case.

there are no solutions of the set (2.41) with L not equal to zero. By combining $-C$ times the second equation with B times the third, and making use of (2.42), we find:

$$(2.44)\quad Lx'[A(B_z - C_y) + B(C_x - A_z) + C(A_y - B_x)] = 0.$$

A similar process gives two other equations, with y' and z' in place of x'. Since L is not zero, and the three first derivatives can not all be zero, by (2.43), we see that the factor involving the coefficients must vanish. But this is the condition of integrability for the Pfaffian, in which case there is an integrating factor M which makes it exact.

On multiplying the first equation of (2.41) by this factor M, and using (2.42) we find:

$$(2.45)\quad A(ML)' + Ly'[(MA)_y - (MB)_x] - Lz'[(MC)_x - (MA)_z = 0.$$

From the fact that M is an integrating factor, the last two terms drop out, and either $(ML)' = 0$, which is to be looked on as a trivial solution or from this and the equations like it. $A = B = C = 0$, and we are at a singular point of the Pfaffian. Thus our assumption that the multiplier of ds in equation (2.3) could be taken as unity was justified.

Returning to the set of equations (2.4), we must impose in addition the relations (2.42) and (2.43). In view of these last two, the first set are not all independent. In fact, from (2.43) we have

$$(2.46)\qquad x'x'' + y'y'' + z'z'' = 0,$$

and on using this and (2.42), we derive the relation:

$$(2.47)\qquad Px' + Qy' + Rz' = 0.$$

Thus the equations to be satisfied reduce to two of the set (2.4), or two independent combinations of these three, supplemented by the conditions (2.42) and (2.43). In other words, we have four differential equations in the four independent

variables x, y, z, L, of the second order in two of these and of the first order in the other two, so that there are apparently six constants of integration. Actually, there are only five constants, which are fixed by the initial and final values of x, y, z, corresponding to the fixed end points and giving the four conditions:

$$(2.5) \qquad \begin{aligned} y_1 &= y(x_1), \qquad z_1 = z(x_1); \\ y_2 &= y(x_2), \qquad z_2 = z(x_2); \end{aligned}$$

together with the initial value of s, which may be taken arbitrarily, and gives the fifth condition:

$$(2.51) \qquad s_1 = s(x_1).$$

That there are only five constants follows from the fact that the system is equivalent to another system of the same order in x, y, z but not containing the derivative of L. This will appear in detail when we come to study local properties of the solutions. The condition (2.51) is of no geometric significance, so that there are a fourfold infinity of curves of shortest length for the general Pfaffian, a double infinity passing through each point.

3. *Some Particular Examples, First Degree Coefficients.* It will throw light on the general situation if we give the calculation of lines of shortest length a little more explicitly for some particular examples. We have shown elsewhere[9] that when the coefficients of the Pfaffian are linear expressions in the variables, the Pfaffian may be reduced to certain simple canonical forms. We omit the cases which are exact differentials, in which case the integral surfaces are families of quadrics, and the lines of shortest length become the geodesics on these quadrics. The lines of shortest length will be unchanged by Euclidean transformations, and under such transformations, the non-integral Pfaffians with linear coefficients may be reduced to four types, namely:

[9] L. c., (2) I, p. 31, 32.

$$axdx + bydy + czdz + C(xdy - ydx) = 0, \tag{3.11}$$

$$axdx + bydy + edz + C(xdy - ydx) = 0, \tag{3.12}$$

$$(ax - by)(adx + bdy) + k(adx - bdy) + czdz = 0, \tag{3.13}$$

$$(ax - by)(adx + bdy) + k(adx - bdy) + edz = 0. \tag{3.14}$$

When we insert the coefficients of these equations in the set (2.4), we have as the first two equations:

$$\begin{aligned} x'' + (ax - Cy)L' - 2CLy' &= 0, \\ y'' + (by + Cx)L' + 2CLx' &= 0, \end{aligned} \tag{3.15}$$

for the types (3.11) and (3.12).

For the last two types, (3.13) and (3.14), we have:

$$\begin{aligned} x'' + (ax - by + k)aL' - 2abLy' &= 0, \\ y'' + (ax - by - k)bL' + 2abLx' &= 0. \end{aligned} \tag{3.16}$$

The third equation of the set (2.4) is

$$z'' + czL' = 0, \tag{3.17}$$

for types (3.11) and (3.13), and

$$z'' + eL' = 0, \tag{3.18}$$

for types (3.12) and (3.14).

The pair of equations (3.15) have no obvious first integral in general, but when we have $a = b$, we may deduce from (3.15):

$$xy'' - yx'' + C(x^2 + y^2)L' + 2CL(xx' + yy') = 0, \tag{3.21}$$

which leads to the integral:

$$xy' - yx' + CL(x^2 + y^2) = K_1. \tag{3.22}$$

We may combine the equations (3.16) so as to give:

$$bx'' + ay'' + 2(ax - by)abL' + 2(ax' - by')abL = 0, \tag{3.23}$$

and integrate this into:

$$bx' + ay' + 2abL(ax - by) = K_2. \tag{3.24}$$

Finally, we note that equation (3.18) integrates directly into:

$$z' + eL = K_3. \tag{3.25}$$

These first integrals enable us to reduce the problem of finding the lines of shortest length to quadratures in two cases, namely type (3.12) with $a = b$, and type (3.14). We shall consider these in order.

We may rewrite the Pfaffian (3.12), with $a = b$. as

$$a(xx' + yy') + ez' + C(xy' - yx') = 0, \tag{3.3}$$

which is to be combined with the identity

$$x'^2 + y'^2 + z'^2 = 1, \tag{3.31}$$

and the integrals (3.22) and (3.25). We may eliminate L from these two integrals, and write the result in the form:

$$e(xy' - yx')/r - Crz' = -CK_3r + eK_1/r, \tag{3.32}$$

where we have set

$$r^2 = x^2 + y^2. \tag{3.33}$$

With this notation, we may replace (3.3) by

$$C(xy' - yx') + ez' = -arr'. \tag{3.34}$$

But:

$$\frac{(xx' + yy')\sqrt{(e^2 + C^2r^2)}}{r} = r'\sqrt{(e^2 + C^2r^2)}, \tag{3.35}$$

is an identity, and if we square both sides of the three equations (3.32), (3.34)′ and (3.35) and add, after making use of (3.31) we find:

$$(1 - r'^2)(e^2 + C^2r^2) = a^2r^2r'^2 + (-CK_3r + eK_1/r)^2. \tag{3.36}$$

From this the relation between r and s may be found by a quadrature.

We next write

$$xy' - yx' = r^2\theta', \tag{3.37}$$

and eliminate z' from equations (3.32) and (3.34), obtaining:

$$r^2\theta'(e^2 + C^2r^2) = e^2K_1 - CK_3er^2 - aCr^2r'. \tag{3.38}$$

This determines θ in terms of r by quadratures, after ds is eliminated by the use of equation (3.36). Finally the z r relation is found from (3.34). We note incidentally that the integrands in r involve the square root of a sixth degree expression in r, but when r^2 is taken as a new variable the integrals are elliptic.

We next consider Pfaffians of type (3.14), which is to be combined with the identity (3.31), and the two integrals (3.24) and (3.25). We eliminate L from these two integrals, and write the result as

$$e(bx' + ay') - 2abuz' = eK_2 - 2abuK_3, \tag{3.41}$$

where we have put

$$u = ax - by. \tag{3.42}$$

With this abbreviation, we may rewrite the Pfaffian (3.14) as

$$\frac{2abu}{c}(bx' + ay') + cez' = -u'\left(kc + \frac{u}{c}[a^2 - b^2]\right), \tag{3.43}$$

where

$$c^2 = a^2 + b^2. \tag{3.44}$$

Finally, we note that

$$(ax' - by')\sqrt{(e^2c^2 + 4a^2b^2u^2)}/c = u'\sqrt{(e^2c^2 + 4a^2b^2u^2)}/c, \tag{3.45}$$

is an identity, and by squaring both sides of the three equations (3.41), (3.43) and (3.45), and then adding, after using the identity (3.31) we deduce:

$$(1 - u'^2/c^2)(e^2c^2 + 4a^2b^2u^2) = (eK_2 - 2abuK_3)^2 + u'^2(kc + [a^2 - b^2]u/c)^2. \tag{3.46}$$

From this the relation of u to s is found by a quadrature.

We next eliminate z' from equations (3.41) and (3.43). The result is:

$$(3.47)\quad (bx' + ay')(e^2c^2 + 4a^2b^2u^2) = \\ -2uu'ab(kc^2 + u[a^2 - b^2]) + ec^2(eK_2 - 2abuK_3).$$

After ds is eliminated from this by use of equation (3.46), a quadrature determines the relation of $(bx + ay)$ to u. Finally, we may find the z u relation from equation (3.41). In this problem all the integrals in u which appear are elliptic integrals.

The two special examples which we have just discussed illustrate the way in which the five constants of integration appear. In each case we had two constants in the two first integrals, and the remaining three constants enter in connection with the three quadratures. One of these constants corresponds to the value of s_1 assigned to the initial point of our arc. But the initial and final points are prescribed. Consequently, in the first case, (3.3), the other four constants are fixed by the requirement that θ_1 and z_1 correspond to r_1, while θ_2 and z_2 correspond to r_2. In the second case (3.14), the four constants are found from the condition that $(bx_1 + ay_1)$ and z_1 correspond to u_1, while $(bx_2 + ay_2)$ and z_2 correspond to u_2. This is in accord with (2.5).

4. *Alternative Procedure for Pfaffians Invariant under a One-parameter Group of Motions.* When we are dealing with a Pfaffian which is invariant under a one-parameter group of rigid motions of space,[10] the procedure used in section 2 may be somewhat simplified. Suppose, first, that the transformation generating the group is a translation, and choose co-ordinates in such a way that the translations of the group are parallel to the z-axis. The Pfaffian may then be written:

$$(4.1)\qquad dz = Mdx + Ndy,$$

in which the coefficients are functions of x and y only. For curves which lie in the Pfaffian, we may calculate the arc length

[10] Cf. l. c., (3) II, p. 101.

from the expression:

$$ds^2 = dx^2 + dy^2 + (Mdx + Ndy)^2. \tag{4.11}$$

Thus, the curves of shortest length will minimize the integral of this expression. But, since we must keep the end points of the arc in space fixed, as well as their projections in the xy plane, it is necessary to restrict the variations by the condition:

$$z_2 - z_1 = \int_{p_1}^{p_2} Mdx + Ndy. \tag{4.12}$$

The application of Lagrange's method of multipliers to the problem at hand leads to the minimization of the integral:

$$I = \int \{\sqrt{x'^2 + y'^2 + (Mx' + Ny')^2} + K(Mx' + Ny')\}dt, \tag{4.13}$$

where the multiplier K is a constant. When we vary the curve, but keep the end points fixed, the first variation of this integral is:

$$\delta I = \int (P\delta x + Q\delta y)dt, \tag{4.14}$$

in which

$$\begin{aligned} P &= d/dt([x' + M(Mx' + Ny')]/s' + KM) \\ &\quad - ([Mx' + Ny']/s' + K)(M_x x' + N_x y'), \\ Q &= d/dt([y' + N(Mx' + Ny')]/s' + KN) \\ &\quad - ([Mx' + Ny']/s' + K)(M_y x' + N_y y'). \end{aligned} \tag{4.15}$$

As before, we take s as our parameter, and change the meaning of the primes so that from now on they denote s derivatives. The vanishing of the first variation of I, which is equivalent to the vanishing of P and Q, accordingly gives the conditions:

$$\begin{aligned} P &= x'' + M(Mx' + Ny')' \\ &\quad + (Mx' + Ny' + K)(M_y - N_x)y' = 0, \\ Q &= y'' + N(Mx' + Ny')' \\ &\quad + (Mx' + Ny' + K)(N_x - M_y)x' = 0. \end{aligned} \tag{4.16}$$

These must be supplemented by the relation

$$x'^2 + y'^2 + (Mx' + Ny')^2 = 1, \tag{4.17}$$

which follows from our choice of s as the parameter, and the Pfaffian, which we rewrite as

$$z' = Mx' + Ny'. \tag{4.18}$$

If we like, we may merely use the last relation to find z in terms of the other variables, after equations (4.16) and (4.17) are solved for x and y in terms of s.

It is interesting to derive this form of the conditions for a line of shortest length from those previously found. Equation (4.17) follows from (2.43) in view of (4.18). For the special form of Pfaffian here considered, we are to put

$$A = M(x, y), \qquad B = N(x, y), \qquad C = -1 \tag{4.2}$$

in the conditions (2.4). This makes two of the three parentheses in the partial derivatives vanish, and in particular the third equation simplifies to:

$$z'' - L' = 0, \tag{4.21}$$

which integrates into

$$z' - L + K = 0. \tag{4.22}$$

From these equations, and equation (4.18), we have:

$$\begin{aligned} L &= Mx' + Ny' + K, \\ L' &= (Mx' + Ny')', \end{aligned} \tag{4.23}$$

and when these are used to eliminate L and L' from the first two of equations (2.4), they reduce to the pair of equations (4.16).

If we regard equation (4.11) as an expression for the arc element on a surface for which $x = \text{const.}$ and $y = \text{const.}$ are the parametric curves, our problem is to find the shortest line through two points on this surface, for which the integral in (4.12) has a pre-assigned value. This is a classical problem of

differential geometry, and the condition may be neatly expressed in terms of the geodesic curvature of the curve.[11]

When the transformation which generates the group of motions under which our Pfaffian is invariant is a helical motion, a set of curvilinear co-ordinates may be introduced such that, when the Pfaffian is expressed in these co-ordinates, one of them does not appear in the coefficients.[12] Using these co-ordinates, we may derive a second form for the differential equations of the lines of shortest length by a procedure quite analogous to that just described.

5. *A General Alternative Procedure.* If we write

$$ds^2 = dx^2 + dy^2 + dz^2 + 2U \cdot V dt^2, \tag{5.1}$$

where

$$U = Ax' + By' + Cz' \tag{5.12}$$

is zero for elements in the Pfaffian, while

$$V = Rx' + Sy' + Tz' \tag{5.13}$$

is an arbitrary linear expression in the derivatives, with coefficients functions of the co-ordinates, we shall have a homogeneous quadratic form in the differentials, independent of the parameter, which may be used to calculate the arc length of curves lying in the Pfaffian. This shows that the lines of shortest length in the Pfaffian may be found by minimizing

$$I = \int \{\sqrt{x'^2 + y'^2 + z'^2 + 2UV} + LU\} dt, \tag{5.14}$$

where L, the multiplier, is a function of t.

The variations are found as before, and give for the first condition:

$$d/dt([x' + \partial(UV)/\partial x']/s' + AL) - [\partial(UV)/\partial x]/s' - L\partial U/\partial x = 0. \tag{5.15}$$

[11] G. Darboux, *Leçons des Surfaces*, Paris, vol. 3, pp. 146–151, § 650, 651, and in particular equation (22).

[12] L. c., (3) II, p. 101.

When the s of (5.1) is taken as the parameter, and the meaning of the primes changed to refer to this s, the condition becomes:

$$x'' + d/ds[\partial(UV)/\partial x'] + (AL)' - \partial(UV)/\partial x - L\partial U/\partial x = 0. \tag{5.16}$$

This equation, and the other two similar to it are not independent. For, on multiplying them respectively by x', y', and z' and adding, we have

$$\Sigma x'x'' + (UV)' + UL' = 0, \tag{5.17}$$

after performing an integration by parts, and also using the fact that

$$\Sigma x'\partial(UV)/\partial x' = 2UV, \tag{5.18}$$

by Euler's theorem on homogeneous functions.

The last term vanishes from the Pfaffian condition, while the first two combined vanish since our choice of parameter makes

$$\Sigma x'^2 + 2UV = 1. \tag{5.19}$$

Thus our conditions are not independent in view of the Pfaffian condition and (5.19). The Pfaffian condition may be used to simplify equation (5.16). For when (UV) is differentiated, the terms still containing U disappear, while those obtained by differentiating U may be combined with those in L to give:

$$x'' + [A(L + V)]' - (L + V)\partial U/\partial x = 0. \tag{5.2}$$

This equation is similar to (2.33), except that $(L + V)$ replaces L in the earlier form.

The process of section 4 is now somewhat clearer. The elimination of z' from the arc element amounted to using a suitable V in (5.14). Sinze z did not explicitly appear in (5.14), while z' appeared just once, with coefficient negative unity in U, the third condition of the form (5.15) reduced to

$$dL/dt = 0, \tag{5.21}$$

so that L reduced to a constant, K.

The V of the present section is

$$\tfrac{1}{2}[z' + (Mx' + Ny')] = V, \tag{5.22}$$

so that, since $U = 0$, we have

$$V = z', \tag{5.23}$$

and

$$L + V = K + z'. \tag{5.24}$$

There may well be other types of Pfaffians where a judicious choice of V would simplify the computation.

6. *The Relation of Lines of Shortest Length to Self Parallel Lines.* Our conditions for a curve of shortest length may be written

$$x'' + AL' + L\dot{y}'(A_y - B_x) - Lz'(C_x - A_z) = 0, \tag{6.1}$$

and the two similar equations of (2.4). These may be written in vector notation, as

$$\mathbf{k} + L'\mathbf{N} - L(\mathbf{t} \times \operatorname{curl} \mathbf{N}) = 0, \tag{6.11}$$

where $\mathbf{k}$ is the curvature vector, x'', y'', z'', $\mathbf{N}$ is the vector normal to the Pfaffian A, B, C and $\mathbf{t}$ is the unit tangent vector, x', y', z'. Thus curl $\mathbf{N}$ is $(C_y - B_z)$, $(A_z - C_x)$, $(B_x - A_y)$.

Any straight line which lies in the Pfaffian is evidently a line of shortest length, since it is a line of absolute minimum length for the three space. For such a line the multiplier L is zero.

Again, if the Pfaffian is integrable, our lines of shortest length should reduce to the geodesics of classical differential geometry for the integral surfaces. Since our Pfaffian condition is not changed by multiplying by any non-vanishing factor,[13] we may in particular use the integrating factor and reduce the integrable case to the exact case, for which curl $\mathbf{N}$ vanishes, and the conditions (6.11) express the well known property of the coincidence of the principal normal of a geodesic on a surface and the normal to the surface, here $\mathbf{N}$.

[13] The multiplier L is divided by this factor, cf. (2.3).

In the integrable case, if we disregard curves through singular lines, in general two points can only be joined by a curve in the Pfaffian if they are on the same surface. If we tried to find the curves in a non-integrable Pfaffian whose principal normals coincided with the normals to the Pfaffian, hereafter referred to as self-parallel curves, we would similarly find that too many conditions were imposed to enable us to draw such a curve through two points in a general relation to one another. However, there will be a single infinity of such curves through each point, one for each direction. Thus each point can be joined with a certain double infinity of points, instead of with all those in three space by a self-parallel curve. In general, these curves will *not* be lines of shortest length for the Pfaffian, and in fact we shall prove that:

THEOREM I. *If a line of shortest length, and a self-parallel line have contact of the second order at a point where the common curvature is not zero, the Pfaffian is integrable at this point.*

For, call $\mathbf{k}$ and $\mathbf{t}$ the common curvature and tangent vectors. From the self-parallel property, $\mathbf{k}$ is parallel to $\mathbf{N}$, and from the condition (6.11) for lines of shortest length, it follows that these two vectors are also parallel to $(\mathbf{t} \times \text{curl } \mathbf{N})$. But, as this last vector is obviously perpendicular to curl $\mathbf{N}$, we have $\mathbf{N}$ perpendicular to curl $\mathbf{N}$, and:

$$(5.12) \qquad \mathbf{N} \cdot \text{curl } \mathbf{N} = \begin{vmatrix} A & B & C \\ \partial/\partial x & \partial/\partial y & \partial/\partial z \\ A & B & C \end{vmatrix} = 0,$$

which is the condition that the Pfaffian be integrable.[14]

We may also formulate the weaker, but more interesting theorem:

THEOREM II. *If a set of curves are simultaneously lines of shortest length, and self-parallel lines, and if through each point of space at least one of these curves passes, having non-zero curvature at the point, the Pfaffian is everywhere integrable.*

[14] Cf. Goursat, E., Lecons sur le probleme de Pfaff, Paris, 1922, p. 24.

We have already noted that for integrable Pfaffians the two types of curves always agree, and also observe now that any straight lines in the Pfaffian are simultaneously lines of shortest length and self-parallel lines. In view of our theorems, these are the only examples of curves in both sets.

7. *Local Properties of Lines of Shortest Length.* To discuss the properties of the lines of shortest length of a Pfaffian in the neighborhood of a point, we require the first few terms in the series solutions of our differential equations. The equations (2.4) may be rewritten as

$$
\begin{aligned}
(7.1)\qquad x'' &= -AL' - Ly'(A_y - B_x) + Lz'(C_x - A_z),\\
y'' &= -BL' - Lz'(B_z - C_y) + Lx'(A_y - B_x),\\
z'' &= -CL' - Lx'(C_x - A_z) + Ly'(B_z - C_y).
\end{aligned}
$$

We have the added conditions, (2.42) and (2.43):

$$(7.11)\qquad Ax' + By' + Cz' = 0,$$

$$(7.12)\qquad x'^2 + y'^2 + z'^2 = 1.$$

We may also make the coefficients satisfy

$$(7.13)\qquad A^2 + B^2 + C^2 = 1,$$

by dividing the Pfaffian by a suitable factor, and shall assume that this has been done in order to simplify the calculations.

If the initial value of any one of the derivatives x', y', z', or a single function of them is given, the three values will then be determined by (7.11) and (7.12). Thus one direction parameter is necessary to characterize a particular line of shortest length. But in addition, we may prescribe the initial value of the multiplier L. We shall now show that this, in general, uniquely fixes the solution.

We begin by differentiating the Pfaffian relation (7.11), to give:

$$(7.14)\qquad \Sigma Ax'' + \Sigma x'A' = 0.$$

Again, from (7.1) and (7.13), we have:

$$(7.15)\quad \Sigma Ax'' = -L' - \Sigma Lx'(BB_x + CC_x - BA_y - CA_z).$$

But, by differentiating (7.13), we find:

$$(7.16)\qquad \Sigma(AA_xx' + BB_xx' + CC_xx') = \Sigma AA' = 0,$$

which may be used to simplify the bracket in (7.15). We make this simplification, solve (7.15) for L', and eliminate the present left member by (7.14). which shows that:

$$(7.17)\quad \Sigma Ax'' = -\Sigma x'A' = -\Sigma\Sigma[A_xx'^2 + (A_y + B_x)x'y' + \cdots].$$

The result is:

$$(7.18)\quad L' = \Sigma\Sigma[A_xx'^2 + (A_y + B_x)x'y' + \cdots] + \Sigma Lx'(AA_x + BA_y + CA_z).$$

This equation gives the initial value of L' in terms of the initial constants described above, namely x', y', z' and L. By differentiating equation (7.18), and using equations (7.1) as they stand, the quantities x'', y'', z'' and L'' are obtained. We have now merely to differentiate the four equations just used successively in order to determine step by step the coefficients in the expansions of x, y, z, L in powers of s, for the neighborhood of any point.

The form of the series just found proves:

THEOREM III. *In general, through each direction in the tangent plane to a Pfaffian at a given point, there are a single infinity of lines of shortest length.*

As we wish to study the curvature more closely, we shall follow out this calculation a little further. However, we shall shorten the work by specializing our axes. We take the origin at the point to be studied, the z axis along the normal to the Pfaffian and the x and y axes for the present any two perpendicular directions in the tangent plane. Consequently, we have initially:

$$(7.19)\qquad A_0 = B_0 = 0, \qquad C_0 = 1, \qquad z_0' = 0.$$

In view of these conditions, when $s = 0$, we have in place of (7.1) and (7.18), the simpler set:

$$(7.2)\quad x_0'' = -L_0 y_0'(A_y - B_x)_0,$$
$$y_0'' = L_0 x_0'(A_y - B_x)_0,$$
$$z_0'' = -L_0' - L_0 x_0'(C_x - A_z)_0 + L_0 y_0'(B_z - C_y)_0.$$

$$(7.21)\quad L_0' = [A_x x'^2 + (A_y + B_x)x'y' + B_y y'^2 + Lx'A_z + Ly'B_z]_0.$$

These conditions also simplify equation (7.16), which becomes:

$$(7.22)\qquad [C_x x' + C_y y']_0 = 0.$$

The last three equations written may be combined so as to give:

$$(7.23)\qquad z_0'' = -[A_x x'^2 + (A_y + B_x)x'y' + B_y y'^2]_0.$$

We may regard the first two of equations (7.2), and (7.23) as determining the curvature vector in terms of our initial conditions, and easily see from them:

THEOREM IV. *The single infinity of lines of shortest length drawn through a common initial point and tangent direction have the same component of curvature normal to the Pfaffian. The component of curvature in the plane of the Pfaffian has a fixed direction, i.e. perpendicular to the tangent, and is in length proportional to the multiplier L. Consequently its length may have any value.*

The same equations also enable us to find the locus of the extremities of the curvature vectors, drawn at a given point, for all the lines of shortest length through that point. This locus is a ruled, cubic surface, whose equation is:

$$(7.24)\quad z(x^2 + y^2) = -A_{x0}y^2 + (A_y + B_x)_0 xy - B_{y0}x^2.$$

We can simplify this equation by a further specialization of axes, which will make

$$(7.25)\qquad (A_y + B_x)_0 = 0.$$

We have merely to take our x and y axes, which have so far only been required to lie in the tangent plane to the Pfaffian, as the

principal directions of the conic obtained by equating the right member of (7.24) to a constant. We note, either from the form and meaning of equation (7.24), or from equation (7.23), that these directions are those of lines of shortest length through the point having greatest and least normal component of curvature respectively.

Since we have taken A, B, C as components of a unit vector along the normal to the Pfaffian, the normal component of curvature is $\Sigma A x''$, which has been evaluated for unspecialized axes in (7.17). Recalling that the tangent to our curves must lie in the tangent plane to the Pfaffian, we see that the special directions are the principal directions of the conic given by:

$$Ax' + By' + Cz' = 0, \tag{7.26}$$
$$\Sigma\Sigma[A_x x'^2 + (A_y + B_x)x'y' \cdots] = 1.$$

The equations satisfied by these principal directions are

$$Ax' + By' + Cz' = 0, \tag{7.27}$$
$$\Sigma[2A_x x' + (A_y + B_z)y' + (A_z + C_x)z'][Cy' - Bz'] = 0.$$

These are the differential equations of a family of lines two of which pass through each point. They are analogous to lines of curvature for surfaces, to which they reduce in the integrable case.

Returning to the specialized axes, and the cubic (7.24), we note that it may reduce to a plane. In fact, this will happen whenever

$$(A_z - B_y)_0 = 0, \tag{7.28}$$

holds simultaneously with (7.25). In particular, if together with (7.25), we have:

$$A_{z0} = B_{y0} = 0, \tag{7.29}$$

the cubic reduces to the tangent plane to the Pfaffian. At such points, the geodesics would have a closer contact with the tangent plane to the Pfaffian than at general points.

For a given direction, the curvature vector is of minimum

length when its component in the tangent plane is zero, that is, when L vanishes. If a line of shortest length had its curvature a minimum for its direction at each of its points, L would be zero throughout, and the curve would be a straight line. In particular, the lines of curvature mentioned above can only be lines of shortest length for non-integrable Pfaffians, when they are straight lines.

We note finally that the equations (7.2) and (7.21) are in accord with our theorem on points at which the condition of integrability is satisfied. For, whenever the principal normal to the curve coincides with the normal to the Pfaffian, the first two curvature components of (7.2) must vanish. Neglecting the case then $L_0 = 0$, in which case the curvature is zero, this requires that:

$$(A_y - B_x)_0 = 0. \tag{7.3}$$

But this is precisely what the condition of integrability,

$$A(C_y - B_z) + B(A_z - B_x) + C(B_x - A_y) = 0, \tag{7.31}$$

reduces to in virtue of the conditions (7.19).

8. *Possible Definitions of Geodesics.* We have defined curves of shortest length on the basis of a minimum property, and have given several methods of finding such curves. We have also mentioned self-parallel curves, although so far we have only defined them by a property of the curvature vector. We shall now show that these curves also arise from a certain variational process, and shall investigate different ways of setting up a variational problem for a class of curves defining types of geodesics for a Pfaffian.

Our lines of shortest length were obtained by minimizing the length of arc joining two points, subject to the condition that the varied curves were situated in the Pfaffian. Had we placed no restrictions on the variations whatever, we would of course have found the equations of geodesics in the Euclidean space, i.e. the straight lines. These would not in general lie in the

Pfaffian. In exceptional cases, for example when the coefficients of the Pfaffian are of the first degree,[9] there are straight lines in the Pfaffian. Since these lines continue to have minimum length when the variations are restricted in any way, they would belong to all the types of geodesics arising from any variational process.

Another possibility is to start with a curve in the Pfaffian, and minimize its length when compared with new curves obtained from it by variations which lie in the Pfaffian.[4] As there are some paradoxical questions connected with this definition, we give the procedure in full. When the end points are fixed, the variation of

$$I = \int ds = \int \sqrt{x'^2 + y'^2 + z'^2}\,dt \tag{8.1}$$

is

$$\delta I = \int [(x'/s')'\delta x + (y'/s')'\delta y + (z'/s')'\delta z]dt, \tag{8.11}$$

or

$$\delta I = \int (x''\delta x + y''\delta y + z''\delta z)ds, \tag{8.12}$$

if s is taken as the parameter, and the primes denote s-differentiation. If this integrand is required to be zero whenever the variations lie in the Pfaffian, i.e. when

$$A\delta x + B\delta y + C\delta z = 0, \tag{8.13}$$

the coefficients of δx, δy, δz must be proportional, and on calling—M the factor of proportionality, we may write the conditions in the form:

$$\begin{aligned} x'' + AM &= 0, \\ y'' + BM &= 0, \\ z'' + CM &= 0. \end{aligned} \tag{8.14}$$

These equations express the fact that the curvature vector of the curve is along the normal to the Pfaffian, which is the property we used before to define the self-parallel curves. We have already shown that, except when these curves are straight lines (for which the property is satisfied in a limiting sense) or curves

at all of whose points the Pfaffian is integrable, they are not lines of shortest length. The form of our conditions (8.14) might have been anticipated by noting that, through the Pfaffian elements of points on a curve even in a non-integrable Pfaffian, a cylindrical strip can be passed. More precisely, a cylindrical surface can be found containing the curve such that the tangent plane to the surface at each point of the curve coincides with the tangent plane to the Pfaffian. The self-parallel curves are the curves which are geodesics on such cylindrical surfaces formed for them. The self-parallel curves also have an interpretation from the standpoint of dynamics, being the orbits of particles moving under no forces in the Pfaffian, i.e. subject to the corresponding non-holonomic constraint.[15] For this reason they have been called mechanical geodesics, in contradistinction to the term geometrical geodesics for the lines of shortest length.[7]

The nature of these two types of geodesics, lines of shortest length and self-parallel lines, is revealed to some extent by their numbers. In a non-integrable Pfaffian there are, as we have seen, a double infinity of lines of shortest length through each point, but only a single infinity of self-parallels. On a surface the two types coincide, there being a single infinity through each point. The lines of shortest length may be looked on as the equilibrium position of a short, tightly stretched string constrained to lie in the configuration. Pairs of endpoints lead to a double infinity through each point for the Pfaffian, and a single infinity for a surface. The self-parallel lines may be looked on as the orbits of particles, under no forces other than those necessary to constrain them to the Pfaffian configuration. As there is one such orbit for each velocity, there are a single infinity through each point for the Pfaffian, as well as in a surface.

9. *Self-parallels for Pfaffians with First Degree Coefficients.* We shall now set up the equations for the self-parallels of Pfaffians with first degree coefficients. As these curves will be unaffected

[15] Cf. Whittaker, E. T., Analytical Dynamics, Cambridge, 1904, p. 245, also (5).

by Euclidean transformations, we may restrict ourselves to the canonical forms previously written in (3.11) to (3.14), i.e.

$$axdx + bydy + czdz + C(xdy - ydx) = 0, \tag{9.11}$$

$$axdx + bydy + edz + C(xdy - ydx) = 0, \tag{9.12}$$

$$(ax - by)(adx + bdy) + k(adx - bdy) + czdz = 0, \tag{9.13}$$

$$(ax - by)(adx + bdy) + k(adx - bdy) + edz = 0, \tag{9.14}$$

which comprise all the non-integrable cases.

When we insert the coefficients of these equations in (8.14), we have as the first two equations:

$$\begin{aligned} x'' + (ax - Cy)M &= 0, \\ y'' + (by + Cx)M &= 0, \end{aligned} \tag{9.15}$$

for the types (9.11) and (9.12).

For the last two types (9.13) and (9.14), we have in place of these:

$$\begin{aligned} x'' + (ax - by + k)aM &= 0, \\ y'' + (ax - by - k)bM &= 0. \end{aligned} \tag{9.16}$$

The third equation of the set (8.14) is

$$z'' + czM = 0, \tag{9.17}$$

for types (9.11) and (9.13), while it is

$$z'' + eM = 0, \tag{9.18}$$

for types (9.12) and (9.14).

For type (9.12) with $a = b$, and type (9.14), we were able to reduce the determination of lines of shortest length to quadratures. For the self-parallel lines, we can reduce the first of these cases to the solution of a single first order equation, while for the second case the self-parallels may be found by quadratures.

In case (9.12) with $a = b$, we have to solve:

$$x'' + (ax - Cy)M = 0, \tag{9.21}$$

$$y'' + (ay + Cx)M = 0, \tag{9.22}$$

$$z'' + eM = 0, \tag{9.23}$$

$$a(xx' + yy') + ez' + C(xy' - yx') = 0, \tag{9.24}$$

$$x'^2 + y'^2 + z'^2 = 1. \tag{9.25}$$

As one of these relations is superfluous, we replace the first three by:

$$e(xy'' - yz'') = C(x^2 + y^2)z'', \tag{9.26}$$

and rewrite the last three equations in polar co-ordinates:

$$arr' + ez' + Cr^2\theta' = 0, \tag{9.27}$$

$$r^2\theta'^2 + r'^2 + z'^2 = 1, \tag{9.28}$$

$$e(r^2\theta')' = Cr^2z''. \tag{9.29}$$

By solving equations (9.27) and (9.28) for z' and θ' in terms of r and r', and combining with the last equation, we obtain a relation between r, r', and r'' which is equivalent to a first order differential equation since s does not appear explicitly. After this r, θ and z may be found in terms of s by quadratures. In the particular case when $a = b = 0$, we have the null system. Here (9.24) and (9.26) combine to show that z'' vanishes, and then it follows from (9.23), (9.21) and (9.22) that M, x'', and y'' vanish, so that the solutions are the straight lines of the null system.

In case (9.14), we have the system:

$$x'' + (ax - by + k)aM = 0, \tag{9.31}$$

$$y'' + (ax - by - k)bM = 0, \tag{9.32}$$

$$z'' + eM = 0, \tag{9.33}$$

$$(ax - by)(ax' + by') + k(ax' - by') + ez' = 0, \tag{9.34}$$

$$x'^2 + y'^2 + z'^2 = 1. \tag{9.35}$$

Here, as before, one equation may be omitted, so we replace the first three equations by

$$e(bx'' - ay'') = 2abkz'', \tag{9.36}$$

which integrates into

$$e(bx' - ay') = 2abkz' + A. \tag{9.37}$$

We next write

$$ax - by = u, \tag{9.38}$$

so that

$$ax' - by' = u', \tag{9.39}$$

and equation (9.34) may be written:

$$u(ax' + by') + ku' + ez' = 0. \tag{9.4}$$

The three equations (9.37), (9.39) and (9.4) may now be solved for x', y' and z' in terms of u and u'. When these values are put in the relation (9.35), an equation in u and u' is obtained, from which the relation of u and s is found by quadratures. After this, three other quadratures determine x, y, z in terms of s.

10. *Pfaffians Invariant under Two Translations.* We have seen that for one of the types of Pfaffians with linear coefficients (3.14), or (9.14), both the lines of shortest length, and the self-parallel lines could be found by quadratures. This is a particular instance of a much wider class of Pfaffians, namely those which, regarded as geometric configurations, may be taken into themselves by two translations. If we take the plane of these translations as the yz plane, the coefficients will depend on x only, and our Pfaffian may be written:

$$A(x)dx + B(x)dy + C(x)dz = 0. \tag{10.1}$$

We shall now determine both kinds of geodesics for this Pfaffian. When s is introduced as the parameter, we have:

$$Ax' + By' + Cz' = 0, \tag{10.11}$$

$$x'^2 + y'^2 + z'^2 = 1, \tag{10.12}$$

as two of the conditions. Since x is a preferred variable, it is natural to solve these for y' and z', obtaining.

$$(10.13)\quad (B^2+C^2)y' = -ABx' + C\sqrt{(B^2+C^2) - (A^2+B^2+C^2)x'^2},$$
$$(B^2+C^2)z' = -ACx' - B\sqrt{(B^2+C^2) - (A^2+B^2+C^2)x'^2}.$$

Let us first investigate the lines of shortest length. In view of the character of the coefficients of (10.1), the conditions (2.4) reduce to:

$$(10.14)\quad \begin{aligned} x'' + AL' - Ly'B_z - Lz'C_z &= 0,\\ y'' + BL' + Lx'B_z &= 0,\\ z'' + CL' + Lx'C_z &= 0. \end{aligned}$$

The last two of these equations are exact, and lead to the integrals:

$$(10.15)\quad \begin{aligned} y' + BL &= c_1,\\ z' + CL &= c_2. \end{aligned}$$

The first equation may be omitted, if (10.11) and (10.12) or the equivalent (10.13) are used. From (10.15), we find:

$$(10.16)\quad Cy' - Bz' = Cc_1 - Bc_2,$$

which may be combined with (10.13) to give:

$$(10.17)\quad (A^2+B^2+C^2)x'^2 = (B^2+C^2) - (Cc_1 - Bc_2)^2.$$

From this equation we may determine the relation of x to s by quadratures, and then use (10.13) to find y and z. Or, from (10.17) and (10.13), y and z may be found as functions of x by quadratures.

We next consider the self-parallel curves. For these, the conditions (8.14) are:

$$(10.2)\quad \begin{aligned} x'' + AM &= 0,\\ y'' + BM &= 0,\\ z'' + CM &= 0. \end{aligned}$$

These may be written:

$$y'' = Bx''/A, \qquad z'' = Cx''/A. \tag{10.21}$$

We now differentiate (10.11) with respect to s, and eliminate y', z', y'' and z'' by (10.13) and (10.21). The result is:

$$\begin{aligned} &x''(A^2 + B^2 + C^2)(B^2 + C^2) \\ &+ x'^2[(B^2 + C^2)AA_x - A^2(BB_x + CC_x)] \\ &+ x'A(CB_x - BC_x)\sqrt{(B^2 + C^2) - (A^2 + B^2 + C^2)x'^2} = 0. \end{aligned} \tag{10.22}$$

If this equation is divided by

$$(B^2+C^2)\sqrt{[A^2+B^2+C^2][(B^2+C^2) - (A^2+B^2+C^2)x'^2]}, \tag{10.23}$$

it becomes exact. To abbreviate the result, and also to aid the geometric interpretation, we denote by t the angle made by the normal to the Pfaffian with the x axis, that is the inclination of the tangent plane to the yz plane. Consequently we have:

$$\sqrt{\frac{A^2 + B^2 + C^2}{B^2 + C^2}} = \csc t, \qquad \frac{A}{\sqrt{A^2 + B^2 + C^2}} = \cos t, \tag{10.24}$$

and the exact equation obtained from (10.22) is:

$$\frac{(x' \csc t)'}{\sqrt{1 - (x' \csc t)^2}} + \cos t \frac{(CB_x - BC_x)x'}{B^2 + C^2} = 0. \tag{10.25}$$

We now put

$$C/B = \tan u, \tag{10.26}$$

so that u is the angle made by the projection of the normal on the yz plane with the y axis. The integral of (10.25) may then be written:

$$\sin^{-1}(x' \csc t) = \int \cos t \, u_x dx + c_3. \tag{10.27}$$

From this, or the equivalent form

$$x' = \sin t \sin \left[\int \cos t\, u_x dx + c_3\right], \tag{10.28}$$

the relation of x to s is found by quadratures, and y and z are found from (10.13) as before.

The appearance of the sine of the integral shows that in general the solutions of equation (10.28) will be totally different from those of equation (10.17). In one case, however, they agree. Namely, when B and C have a constant ratio, both of the equations are equivalent to

$$x' \csc t = \text{const.} \tag{10.29}$$

This is in accord with theorem II, since the Pfaffian (10.1) is integrable if and only if the ratio of B to C is constant.

Let us draw a line X_P, through the point under consideration and parallel to the projection of the x axis on the tangent plane of the Pfaffian. Denote the projection of ds on this line by dx_P. Then, since t is the inclination of the tangent plane to the yz plane, i.e. the complement of the angle between the X and X_P axes, we have

$$dx = dx_P \sin t. \tag{10.3}$$

Thus equation (10.29) expresses the constancy of dx_P/ds, the sine of the inclination of ds to the line Y_P drawn in the tangent plane to the Pfaffian perpendicular to X_P. But, when B/C is constant, the lines Y_P are all parallel to the fixed direction:

$$\frac{dx}{0} = \frac{dy}{B} = \frac{dz}{C}, \tag{10.31}$$

the integral surfaces are cylinders with elements in this direction, and our condition simply states that the geodesics on these cylinders are helices.

In the general case we may draw X_P and Y_P as before. We note that Y_P is parallel to the intersection of the tangent plane to the Pfaffian with the yz plane. As this is perpendicular to the

projection of the normal on this plane, du/ds measures the rate of rotation of Y_P. If v is the inclination of ds to Y_P, dv/ds, measures the rate of rotation of ds about Y_P. With this notation, we may write (10.25) in the form:

$$dv = du \cos t, \tag{10.32}$$

which brings out its geometric significance.

11. *Pfaffians Invariant under Two Rigid Displacements.* For the type of Pfaffian with linear coefficients (3.12) or (9.12), we were able to reduce the determination of the lines of shortest length to quadratures, and that of the self parallel lines to a first order differential equation. This is also typical of a wider class of Pfaffians, namely those which, regarded as geometrical configurations, may be taken into themselves by two rigid displacements. In this case[10] we may regard the two parameter group of motions as generated by a rotation, and a translation parallel to the axis of rotation. We take the axis of rotation as the z axis, and introduce polar co-ordinates. Our Pfaffian may then be written:

$$dz = F(r)rdr + G(r)r^2d\theta, \tag{11.1}$$

or:

$$dz = (xF - yG)dx + (yF + xG)dy. \tag{11.11}$$

In determining either type of geodesic, we have the two relations:

$$z' = Frr' + Gr^2\theta', \tag{11.12}$$

$$r'^2 + r^2\theta'^2 + z'^2 = 1, \tag{11.13}$$

when the arc length is taken as the parameter. These may be solved for θ' and z', to give:

$$r\theta'(r^2G^2 + 1) = - FGr^2r' + \sqrt{r^2G^2 + 1 - r'^2(r^2G^2 + 1 + r^2F^2)}, \tag{11.14}$$

$$z'(r^2G^2 + 1) = Frr' + Gr\sqrt{r^2G^2 + 1 - r'^2(r^2G^2 + 1 + r^2F^2)}. \tag{11.15}$$

When we seek the lines of shortest length, we apply the conditions (2.4) to the Pfaffian (11.11), and have:

$$
\begin{aligned}
(11.16)\qquad & x'' + (xF - yG)L' - Ly'(2G + rG_r) = 0,\\
& y'' + (yF + xG)L' + Lx'(2G + rG_r) = 0,\\
& z'' - L' = 0.
\end{aligned}
$$

The last of these may be integrated at once, into:

$$(11.17)\qquad z' - L = c_1.$$

The first two may be combined so as to give:

$$(11.18)\qquad xy'' - yx'' + r^2GL' + Lrr'(2G + rG_r),$$

which has as its integral:

$$(11.19)\qquad r^2\theta' + r^2GL = c_2.$$

On eliminating L from this by means of (11.17), we obtain:

$$(11.2)\qquad r^2(Gz' + \theta') = c_2 + c_1r^2G.$$

After we eliminate θ' and z' from this equation by means of (11.14) and (11.15), there results a relation between r and r' from which s is found in terms of r by a quadrature. We then find θ and z in terms of s or r by further quadratures.

For the self parallel lines, the conditions (8.14) are:

$$
\begin{aligned}
(11.21)\qquad & x'' + M(xF - yG) = 0,\\
& y'' + M(yF + xG) = 0,\\
& z'' - M = 0.
\end{aligned}
$$

The first two of these equations may be replaced by:

$$
\begin{aligned}
(11.22)\qquad & (r^2\theta')' + Mr^2G = 0,\\
& r'' - r\theta'^2 + MrF = 0,
\end{aligned}
$$

which may be combined with the third equation to give:

$$
\begin{aligned}
(11.23)\qquad & z''rF = r\theta'^2 - r'',\\
& \theta''rF = r''G - r\theta'^2G - 2Fr'\theta'.
\end{aligned}
$$

We now differentiate (11.12) with respect to s, eliminating z'' and θ'' by means of (11.23), and then eliminating θ' from the result by means of (11.14). The result is a relation between r, r' and r'' which does not contain s explicitly, and hence reduces the determination of the dependence of r on s to the solution of a first order equation. The solution is again completed by using (11.14) and (11.15).

We notice that for the lines of shortest length, the relation of r to s involved three constants, namely a constant of integration in addition to the c_1 and c_2 explicitly introduced. For the self parallels, the corresponding relation comes from a second order differential equation with no constants explicitly introduced, and so involves two constants.

12. *Pfaffians in Riemannian Spaces.* Instead of restricting ourselves to the case of a single Pfaffian in Euclidean three space, we may study the more general configuration of m Pfaffians in an n-dimensional Riemannian space.[8] Let its metric be given by:

$$ds^2 = g_{ij}dx_i dx_j, \quad i, j = 1, 2, \cdots n, \tag{12.1}$$

where we use the summation convention.[16] The equations of the m embedded Pfaffians may be written:

$$A_{ai}dx_i = 0, \quad a = 1, 2, \cdots m. \tag{12.11}$$

The *geodesics* in the Riemannian space are defined by

$$\delta \int ds = 0, \tag{12.12}$$

which is equivalent to:

$$\{g_{ij}x_j'' + [jk, i]x_j'x_k'\}\delta x_i = 0. \tag{12.13}$$

Here primes denote s-differentiation.

For the self-parallel curves, we minimize the integral in (12.12), subject to the condition that the variations satisfy the m relations (12.11). As before, we accomplish this by introducing m mul-

[16] Cf. Eisenhart, L. P., *Riemannian Geometry*, Princeton, 1926, pp. 1–17 for the notation used in this paragraph.

tipliers, M^a, to combine (12.11) and (12.13) into a single equation, in which the deltas may then be taken arbitrarily. Thus the equations of the *self parallel* lines are found to be:

$$g_{ij}x_j'' + [jk, i]x_j'x_k' + M^aA_{ai} = 0, \tag{12.14}$$

which must be taken in conjunction with

$$A_{ai}x_i' = 0, \tag{12.15}$$

and

$$g_{ij}x_i'x_j' = 1. \tag{12.16}$$

The set of equations (12.14) is equivalent to the set:

$$x_i'' + \left\{ \begin{matrix} i \\ j\ k \end{matrix} \right\} x_j'x_k' + M^aA_a{}^i = 0. \tag{12.17}$$

For the lines of shortest length, we must minimize the integral in (12.12), so restricting the variations that the varied curves lie in the Pfaffian system (12.11). This is accomplished by introducing variable multipliers L^a, and then minimizing:

$$\int (ds + L^aA_{ai}dx_i). \tag{12.18}$$

This leads to the system of equations:

$$g_{ij}x_j'' + [jk, i]x_j'x_k' + (L^aA_{ai})' - L^a\partial A_{aj}/\partial x_i\, x_j' = 0, \tag{12.19}$$

which, again, must be taken in conjunction with equations (12.15) and (12.16). We may also derive the alternative set to (12.19):

$$x_i'' + \left\{ \begin{matrix} i \\ j\ k \end{matrix} \right\} x_j'x_k' + g^{ih}[(L^aA_{ah})' - L^a\partial A_{aj}/\partial x_h\, x_j']. \tag{12.2}$$

The last two equations written as the conditions for the lines of *shortest length* may be expanded, and in fact simplify somewhat when written more explicitly. They become:

$$g_{ij}x_j'' + [jk, i]x_j'x_k' + L^{a\prime}A_{ai} + L^ax_k'(A_{ai,\,k} - A_{ak,\,i}) = 0, \tag{12.21}$$

and

$$x_i'' + \left\{ \begin{matrix} i \\ j\ k \end{matrix} \right\} x_j' x_k' + L^{a'} A^i{}_a + L^a x_k' (A^i{}_{a,\,k} - A_{ak,}{}^i) = 0, \tag{12.22}$$

respectively. The terms here written as covariant or contravariant derivatives must be formed with respect to the indices i and k only, since the index a is not of tensor character, but merely serves to distinguish the different Pfaffian conditions. In (12.21) the covariant derivatives may be put in place of the corresponding partial derivatives, since we are dealing with a curl for which the added terms cancel out.

A comparison of (12.13) with (12.14) and (12.19) or (12.21) shows that, if any geodesic of the n-space lies in the Pfaffian configuration, it will satisfy the equations for both other types of geodesics, self parallels and lines of shortest length, with zero multipliers in each case, M^a and L^a respectively. In general, there will be no such curves.

If all the Pfaffian relations of the set (12.11) are integrable, at a point, the curl of their coefficients will vanish:

$$A_{ai,\,k} - A_{ak,\,i} = 0, \quad a = 1, 2, \cdots m. \tag{12.23}$$

and the solutions of (12.14) and (12.21) coincide, with

$$M^a = L^{a'}, \tag{12.24}$$

There may be intermediate cases in which the curves of the two types agree because some of the relations are integrable, while the non-integrable ones are compensated for by special properties of the Riemannian space and the curves considered. Or, we may merely have integrable linear combinations, since the problem we are considering is clearly unchanged when the Pfaffians are replaced by any independent linear combinations, except for simple transformations of the multipliers.

SYNTHESIS OF A FINITE TWO-TERMINAL NETWORK WHOSE DRIVING-POINT IMPEDANCE IS A PRESCRIBED FUNCTION OF FREQUENCY

By Otto Brune[1]

Contents

Part I. Introduction 191
1. Statement of the Problem 191
2. Contributions of Previous Investigators 192
3. Necessary Conditions to be satisfied by the Impedance Function of a Passive Network 194
Part II. Properties of "Positive Real" Functions 197
1. Definitions 197
2. Theorems 198
Part III. Synthesis of Networks with Prescribed Impedance Function. 208
1. Functions with not more than two poles in the interior of the left half plane 208
2. General Procedure 218
3. Discussion of networks obtained 225
Part IV. Extension of the Method to the Determination of certain Equivalent Networks 228
1. Change in Order of Procedure 228
2. Consideration of "Imaginary Frequencies." 229
3. Relation to Equivalent Networks obtained by affine transformation of Quadratic Forms 232

Part I. Introduction

1. *Statement of the Problem*

In the well known methods of analysing the performance of linear passive electrical networks with lumped network elements it is usual to derive from the given structure of the network a scalar function $Z(\lambda)$ known as the impedance function of the network; this function determines completely the performance

[1] Containing the principal results of a research submitted for a doctor's degree in the Department of Electrical Engineering, Massachusetts Institute of Technology. The author is indebted to Dr. W. Cauer who suggested this research.

of the network as far as current and voltage at one pair of terminals is concerned.

The converse problem is assuming increasing importance in electrical engineering practice, namely: Given the behavior of the network at its terminals, to find the structure of the network. The solution of this problem is not unique; for the present we shall concern ourselves mainly with finding at least one solution.

The behavior of a network at a pair of terminals may be prescribed in several different ways; we shall assume that it is prescribed by specifying a function $Z(\lambda)$ to be the impedance function of the network, all other methods being reducible to this one.

It is evident that certain restrictions are imposed on the possible forms of $Z(\lambda)$ by the physical conditions inherent in a passive network with lumped linear elements. A concise statement of the necessary and sufficient conditions is an essential part of the solution of the problem.

The problem may therefore be stated as follows:—

To find the necessary and sufficient conditions to be satisfied by the impedance function of a finite passive network, and to construct a network corresponding to any function satisfying these conditions.

2. *Contributions of previous Investigators*

The first important contribution to the theory of synthesis of two-terminal networks was probably made by R. M. Foster when, in his "Reactance Theorem" (1) he gave the necessary and sufficient conditions which must be satisfied by the impedance function of a *purely reactive network*. They are:

The poles and zeros of $Z(\lambda)$ are pure imaginary and mutually separate each other (this includes the necessity for a pole or zero at the origin and at infinity).

Foster gives two methods of constructing a network when the prescribed impedance function satisfies these conditions. These methods correspond to a development of $Z(\lambda)$ or its reciprocal in partial fractions; the networks so obtained are

shown in Fig. 1 (*a*) and (*b*). Cauer (3) has amplified these results by pointing out a continued fraction development corresponding to the equivalent networks shown in Fig. 1 (*c*) and (*d*).

Cauer (3) has also adapted the results for purely reactive networks to all networks containing only two kinds of elements, i.e. such networks as contain only resistance and capacitance or

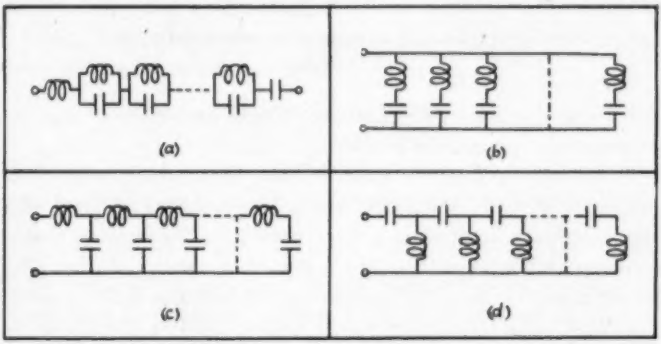


FIG. 1

only resistance and inductance. For such networks the prescribed $Z(\lambda)$ *must have all its poles and zeros on the negative real axis and the poles and zeros must mutually separate each other.* Corresponding networks may be constructed similar to those in Fig. 1 but having either coils or condensers (as the case may be) replaced by resistances.

The particular case of our problem in which only two kinds of elements occur has thus been completely solved.

Both Foster (2) and Cauer (3) have studied the case when $Z(\lambda)$ is capable of representation by a two-mesh network having three kinds of elements. Such a function must have the form

$$Z(\lambda) = \frac{a_0 + a_1\lambda + a_2\lambda^2 + a_3\lambda^3 + a_4\lambda^4}{b_1\lambda + b_2\lambda^2 + b_3\lambda^3},$$

where the values of the coefficients in the numerator and denominator are restricted by certain inequalities (some of the coefficients may of course be zero).

The network shown in Fig. 2 however has an impedance function of the same form as a two-mesh network but with coefficients which do not satisfy the inequalities necessary for the impedance function of a two-mesh network. Consequently the

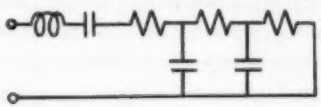


FIG. 2

conditions discussed by Foster and Cauer are sufficient but not necessary in the general problem.

T. C. Fry (4) has discussed in some detail the possibility of using the Stieltjes' continued fraction as a means of developing the impedance function in a form usable for the construction of a corresponding network. He is guided in this by the fact that a continued fraction development of the function is the counterpart of a ladder structure in the network. The conditions for the applicability of a Stieltjes continued fraction expansion of a function are sufficient for the function to be an impedance function of an electric network.[2] Fry also succeeds in extending the method to functions which do not immediately satisfy these conditions by means of several ingenious transformations. The method is not general however, and since it leads to infinite networks in many cases, does not seem to be the natural method of attack.

In all the above cases the conditions imposed on $Z(\lambda)$ are sufficient for the construction of a network but not necessary from the general point of view.

3. *Necessary Conditions to be Satisfied by the Impedance Function of a Finite Network*

Since the impedance function of every finite network is a rational function, it is clear that we must be able to write every

[2] See also Cauer, loc. cit. (3).

prescribed impedance function in the form

$$Z(\lambda) = \frac{f(\lambda)}{g(\lambda)}, \tag{1}$$

where

$$f(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \cdots a_n\lambda^n, \tag{1a}$$

$$g(\lambda) = b_0 + b_1\lambda + b_2\lambda^2 + \cdots b_m\lambda^m. \tag{1b}$$

Now if the network be disturbed it is a well known fact that the transient current is given by

$$I_t = \sum_k C_k \epsilon^{\alpha_k t}, \tag{2}$$

where α_k are the roots of $f(\lambda)$; and since transient currents cannot increase without limit in a passive network, we must have [3]

$$\text{Re } \alpha_k \leq 0. \tag{2a}$$

In the above the terminals of the network were regarded as closed, i.e. the transient currents are such as to cause no voltage drop across the terminals. Similar considerations for open terminals give the condition

$$\text{Re } \beta_k \leq 0, \tag{2b}$$

where β_k are the roots of $g(\lambda) = 0$.

A further necessary condition is obtained by considering a steady alternating voltage $E = E_m \cos \omega t$ applied. We may then write

$$E = E_m \text{ Re } \epsilon^{i\omega t},$$

$$I_s = E_m \text{ Re} \frac{\epsilon^{i\omega t}}{Z(i\omega)}$$

$$= \frac{E_m}{|Z(i\omega)|} \cos(\omega t - \varphi) \quad \text{where} \quad \varphi = \arg Z(i\omega).$$

From the fact that the average power taken by the network cannot be negative, it follows that $|\varphi| \leq \pi/2$ or

$$\text{Re } Z(i\omega) \geq 0. \tag{3}$$

[3] The notation Re means "real part of" throughout.

We may also consider an applied voltage

$$E = \epsilon^{\gamma t} \cos \omega t$$
$$= \text{Re}\, \epsilon^{(\gamma + i\omega)t}.$$

The "steady state" current may be calculated as before as

$$I_s = \text{Re}\, \frac{\epsilon^{(\gamma+i\omega)t}}{Z(\gamma + i\omega)}$$
$$= \frac{\epsilon^{\gamma t}}{|Z(\gamma + i\omega)|} \cos(\omega t + \varphi) \quad \text{where} \quad \varphi = \arg Z(\gamma + i\omega).$$

If γ is *positive* the transient currents can, after a sufficiently long time, be neglected (even if the network is purely reactive). If on the other hand γ is *negative*, the transient term may become more important than the steady state term, and no conclusions concerning the energy taken by the network can be drawn by considering the "steady state" current alone; in fact the application of the term "steady state" is misleading in this case. Assuming γ to be positive, however, let us determine the energy W taken by the network after the transient current has become negligible, and consider $t = 0$ at the beginning of this time, the voltage then being unity. Thus

$$E = \epsilon^{\gamma t} \cos \omega t; \qquad I = \frac{1}{|Z(\gamma + i\omega)|} \epsilon^{\gamma t} \cos(\omega t - \varphi).$$

Integrating the product EI over a time τ

$$W = \frac{1}{|Z(\gamma + i\omega)|} \int_0^\tau \epsilon^{2\gamma t} \cos \omega t \cos(\omega t - \varphi) dt$$
$$= \frac{1}{4|\gamma + i\omega| \cdot |Z(\gamma + i\omega)|} \left\{ \epsilon^{2\gamma\tau} \cos(2\omega\tau - \varphi - \chi) - \cos(\varphi + \chi) + \frac{\cos\varphi}{\cos\chi} (\epsilon^{2\gamma\tau} - 1) \right\}$$
$$\text{where} \qquad \chi = \arg(\gamma + i\omega).$$

It is evident that as τ increases the value of W will depend

entirely on the terms having $\epsilon^{2\gamma\tau}$ as a factor. Consequently if W is never to be negative we must have $\frac{\cos\varphi}{\cos\chi} \geq 1$ (the equality sign holding only for the non-dissipative case). This apparently requires besides

$$\text{Re } Z(\lambda) \geq 0 \quad \text{where Re } \lambda \geq 0; \tag{4a}$$

also

$$|\arg Z(\lambda)| \leq |\arg \lambda| \quad \text{when } |\arg \lambda| \leq \frac{\pi}{2}. \tag{4b}$$

We have enumerated a number of necessary conditions, but it is evident that they are not all independent. It will next be shown that all other conditions are consequences of either (4*a*) or (4*b*). Several other important properties of functions satisfying these conditions will also be noted; these will be of value in finding a network corresponding to a given impedance function.

Part II—Properties of "Positive Real" Functions

1. *Definitions.* For the sake of brevity we define the following terms:

A "*positive function*" $Z(\lambda)$ is a function whose real part is positive when the real part of λ is positive, i.e. a function which satisfies condition (4*a*). If in addition $Z(\lambda)$ is real when λ is real, it will be called a "*positive real*" function. The quotation marks will be retained as a reminder of the special sense in which these terms are being used.

In general the notation $Z(\lambda)$ will imply that the function falls under the above definition. If a function does not or is not known to satisfy these conditions, an asterisk will be added or another notation employed.

If a function is the impedance function of a physical network, it will be said to have a *network representation*. The process of finding such a network will be called finding a network representation of the function.

The proof that every "positive real" function has a network

representation is reserved for treatment in Part III of this paper. In the meantime we shall frequently anticipate this fact in interpreting mathematical properties of " positive real " functions.

The discussions which follow depend for the most part on well known properties of functions of a complex variable. For more ready reference however they are stated in the form of theorems.

2. *Theorems on " Positive Real " Functions*

The idea of a " positive " function can be interpreted very conveniently in terms of the transformation of one complex plane into another. By the equation $Z = Z(\lambda)$ every point in the λ plane determines uniquely a point in the Z plane. The property of " positiveness " as defined by (4*a*) means that if λ is chosen in the right half plane, Z will fall in the right half plane, but not necessarily conversely. Consequently if $Z(\lambda)$ is a " positive " function, the transformation $Z = Z(\lambda)$ will transform the right half of the λ plane into a part of the right half of the Z plane (except in special cases where the right half λ plane

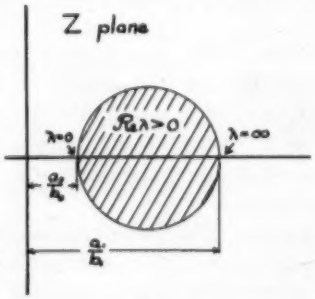


FIG. 3

goes into the whole right half Z plane). A simple example of this is shown in Fig. 3 where $Z = \dfrac{a_0 + a_1\lambda}{b_0 + b_1\lambda}$.

We thus have

THEOREM I. If

$$Z = Z(\lambda)$$

and

$$W' = W(\lambda),$$

Z and W both being "positive functions," then

$$W[Z(\lambda)]$$

will be a "positive" function of λ.

This theorem can obviously be applied in two ways illustrated by the two corollaries:

THEOR. I, COROLL. 1. If $Z(\lambda)$ is a "positive" function, $Y(\lambda) = \frac{1}{Z(\lambda)}$ will be a "positive "function.

THEOR. I, COROLL. 2. If $Z(\lambda)$ is a "positive" function, $Z_1(\lambda) = Z\left(\frac{1}{\lambda}\right)$ will be a "positive" function.

The physical interpretation of Corollary 1 is: to every network corresponds a "reciprocal" network whose admittance characteristic is the impedance characteristic of the first network.

On the other hand Coroll. 2 means that to every network corresponds another whose impedance characteristic is that of the first network but with the reciprocal frequency scale (e.g. frequency is replaced by wave-length).

Taken more generally Theorem I states that if in $Z(\lambda)$, the impedance function of a given network N, λ be replaced by the impedance function $z(\lambda)$ of another network n, the impedance function of a third network N' will result. In certain cases N' can be derived from N by replacing inductances by networks proportional to n, and capacities by networks proportional to the network $(1/n)$, reciprocal to n. If the conception of an ideal transformer be admitted a representation of mutual elements in terms of n is also possible and N' can be constructed directly and completely from N by such substitions.

We next come to a discussion of the poles and zeroes of a "positive" function. We note first of all the

Lemma 1: A zero of multiplicity n is surrounded by $2n$ sectors of equal angles (Fig. 4) in which the real part of the function is

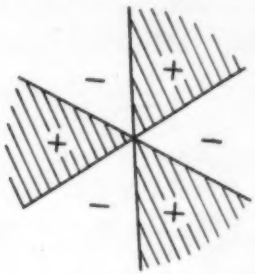


FIG. 4

alternately positive and negative. For at such a point $\lambda = c$ the Taylor's series for the function becomes

$$Z(\lambda) = \left\{ \frac{d^n}{d\lambda^n} \frac{Z(\lambda)}{n!} \right\}_{\lambda=c} (\lambda - c)^n + \text{term of higher degree}$$

In the immediate neighbourhood of the point $\lambda = c$ the first term predominates. Placing

$$(\lambda - c) = \rho\epsilon^{i\theta} \quad \text{and} \quad \left\{ \frac{d^n}{d\lambda^n} \frac{Z(\lambda)}{n!} \right\}_{\lambda=c} = k\epsilon^{i\varphi} \quad (k, \varphi \text{ real constants})$$

we have

$$Z(\lambda) = Z(\rho, \theta) \approx k\rho^n \epsilon^{i(n\theta+\varphi)}$$
$$\text{Re } Z(\lambda) \approx k\rho^n \cos(n\theta + \varphi), \tag{5}$$

whence the truth of the lemma is evident.

A zero within the right half plane would therefore obviously violate the condition of "positiveness." Moreover on the boundary of this region (i.e. on the imaginary axis) only two sectors are possible; the real part of the function must be positive on the one side of the boundary, negative on the other. This means that zeros on the imaginary axis must be simple, and that φ in equation (5) must be zero, i.e. $(dZ/d\lambda)_{\lambda=i\sigma}$ = a real positive constant.

Corresponding conditions for the poles are immediately evident from Theorem I, Coroll. 1. For a pole, however, the condition $\left(\frac{d}{d\lambda}\frac{1}{Z}\right)_{\lambda=i\sigma}$ = a real positive constant, can be expressed more simply by saying that the residue of the function at $\lambda = i\sigma$ must be real and positive. All these facts are summarized in

THEOREM II. If $Z(\lambda)$ is a "positive function" then

(i) its zeros and poles must lie in the left half λ plane or on the boundary;
(ii) zeros and poles on the imaginary λ axis must be simple;
(iii) at a zero on the imaginary λ axis $dZ/d\lambda$ is a real positive constant while at a pole on the boundary the residue is a real positive constant.

Of especial interest are the zeros and poles at $\lambda = 0$ and $\lambda = \infty$. These points lie on the boundary; consequently at them only simple poles and zeros can exist.

If $\lambda = 0$ is a simple zero or a pole, λ will be a factor in the numerator or denominator respectively. If $\lambda = \infty$ is a simple pole or zero, the degree of the numerator or denominator will be greater by one than the degree of the denominator or numerator respectively. This consequence of Theorem II (ii) can therefore be summarized as

THEOREM II, COROLL. 1. The degree in λ or in $1/\lambda$ of numerator and denominator of $Z(\lambda)$ cannot differ by more than 1.

We interpose here a well-known theorem in function theory, which finds special application in the theory of potentials. This theorem states that a function $R(x, y)$ which is regular within a certain region attains its maximum and minimum value on the boundary of that region. If poles occur on the boundary they may be excluded by arbitrarily small semi-circles, the theorem then holding for the modified boundary. In the case of "positive" functions the value of the real part becomes arbitrarily large on such semi-circles if the radius is made arbitrarily small, but remains positive. Hence we have

THEOREM III. If $Z(\lambda)$ is a "positive" function the real part of $Z(\lambda)$ in the right half plane attains its *minimum* value on the axis of imaginaries.

THEOREM III, COROLL. 1. If R_1 is equal to or less than the minimum value of the real part of $Z(i\omega)$ then

$$Z'(\lambda) = Z(\lambda) - R_1$$

is a "positive" function.

We shall apply this theorem immediately to the separation of poles on the axis of imaginaries from the function.

Consider first a pole at $\lambda = \infty$. The numerator is of degree one higher than the denominator, i.e.

$$Z(\lambda) = \frac{a_{n+1}\lambda^{n+1} + a_n\lambda^n + \cdots a_1\lambda + a_0}{b_n\lambda^n + \cdots b_1\lambda + b_0}.$$

By ordinary division of numerator into denominator

$$\begin{aligned} Z(\lambda) &= L\lambda + \frac{a_n'\lambda^n + \cdots a_1'\lambda + a_0'}{b_n\lambda^n + \cdots b_1\lambda + b_0}, \\ &= L\lambda + Z'(\lambda) \end{aligned}$$

where

$$L = \frac{a_{n+1}}{b_n}.$$

Note that the real part of $L\lambda$ on the axis of imaginaries is zero, consequently $\operatorname{Re} Z'(i\omega) = \operatorname{Re} Z(i\omega) \geq 0$. Since $\operatorname{Re} Z'(\lambda)$ is regular in the right half plane however (i.e. has no poles there) it is everywhere within this region greater than the minimum value on the boundary; i.e. > 0. hence $Z'(\lambda)$ is a "positive" function.

The case when $\lambda = 0$ is a pole follows from the above by making the substitution $\lambda' = 1/\lambda$ everywhere. Then

$$\begin{aligned} Z(\lambda) &= \frac{a_0 + a_1\lambda + a_2\lambda^2 + \cdots a_n\lambda^n}{b_1\lambda + b_2\lambda^2 + \cdots b_m\lambda^m} \\ &= \frac{D}{\lambda} + \frac{a_1' + a_2'\lambda + \cdots a_n'\lambda^{n-1}}{b_1 + b_2\lambda + \cdots b_m\lambda^{m-1}} \end{aligned}$$

$$= \frac{D}{\lambda} + Z'(\lambda), \quad \text{where } D = \frac{a_0}{b_1}.$$

and $Z'(\lambda)$ is a " positive real " function.

If $Z(\lambda)$ has a pole on the boundary at $\lambda = i\sigma$ (say) the condition of reality on the real axis necessitates the presence of the conjugate pole. These poles can always be removed in the following way, which is merely an application of partial fractions.

Let

$$Z(\lambda) = \frac{f(\lambda)}{g(\lambda)} = \frac{f(\lambda)}{(\lambda^2 + \sigma^2)g_1(\lambda)}.$$

Then

$$Z(\lambda) = \frac{k}{\lambda + i\sigma} + \frac{k}{\lambda - i\sigma} + \frac{f_1(\lambda)}{g_1(\lambda)}$$

(k being the real positive residue of $Z(\lambda)$ at $\lambda = \pm i\sigma$)

$$Z(\lambda) = \frac{2k\lambda}{\lambda^2 + \sigma^2} + Z'(\lambda). \tag{6}$$

The important difference between equation (6) and the ordinary partial fraction development is that the numerator of the first partial fraction would ordinarily have been $2(k\lambda + h\sigma)$; here however because the residue at $\lambda = i\sigma$ is real, $h = 0$. This fact enables us to conclude that the real part of the first partial fraction in equation (18) is zero on the axis of imaginaries whence it follows as before that $Z'(\lambda)$ is a " positive " function. Summing up we have

THEOREM IV. If $Z(\lambda)$ is a " positive real " function which has poles (residue $= k_r/2$) at $\lambda = i\sigma_0, i\sigma_2, \cdots i\sigma_n, \infty$. Then

$$Z(\lambda) = L\lambda + \sum_{r=0}^{n} \frac{k_r\lambda}{\lambda^2 + \sigma_r^2} + Z'(\lambda), \tag{7}$$

where each term is a " positive real " function and

$$\text{Re } Z'(i\omega) = \text{Re } Z(i\omega).$$

THEOREM IV, COROLL. 1. If $Z(\lambda)$ is a " positive real " func-

tion which has zeros at $\lambda = 0, \cdots i\sigma_r \cdots \infty$ then

$$\frac{1}{Z(\lambda)} = C\lambda + \sum_{r=0}^{n} \frac{k_r\lambda}{\lambda^2 + \sigma_r^2} + \frac{1}{Z'(\lambda)}, \tag{7a}$$

where each term is a " positive real " function and

$$\mathrm{Re}\left[\frac{1}{Z'(i\omega)}\right] = \mathrm{Re}\left[\frac{1}{Z(i\omega)}\right].$$

THEOREM IV. COROLL. 2. If $Z(\lambda)$ is a " positive real " function all of whose poles lie on the imaginary axis then $Z'(\lambda)$ in (7) is a real positive constant. A similar statement holds when all the zeros lie on the axis of imaginaries by applying Corollary 1.

Corresponding to Theorem III above is the well-known fact that a function $R(x, y)$ which is regular within a certain region is completely specified within that region by the values which it takes on the boundary. There is obviously a connection between this fact and the possibility of stating the necessary and sufficient conditions in terms of conditions on the boundary.

We note first of all that the condition of regularity in the right-half plane excludes the possibility of poles within this region, i.e. is synonymous with equation (2*b*). We have seen further that poles on the imaginary axis are independent of the real part of $Z(i\omega)$ (Theorem IV). Consequently the condition to be satisfied by poles on the boundary must be stated separately. If in addition to this we state that the real part of $Z(i\omega)$ is positive, all conditions for a " positive " function are fulfilled. We thus have

THEOREM V. If $Z(\lambda)$ is a function such that

(i) No poles lie in the right half plane.

(ii) Poles on the boundary are simple and have positive real residues.

(iii) $\mathrm{Re}\, Z(i\omega) \geq 0$.

then $Z(\lambda)$ is a " positive " function. In particular the zeros $\lambda = \alpha_r$ will satisfy the condition (2*a*).

By Theorem I, Coroll. 1 we have also

THEOREM V, COROLL. 1. If $Z(\lambda)$ satisfies the conditions

(i) No zeros lie within the right half plane.

(ii) Zeros on the boundary are simple and at them $dZ/d\lambda$ = a positive real constant.

(iii) Re $Z(i\omega) \geq 0$.

then $Z(\lambda)$ is a " positive " function. In particular the poles $\lambda = \beta_r$ will satisfy the condition (*2b*).

It is always possible to reduce a given function to one in which the second restriction in Theor. V is not necessary by the procedure of Theor. IV.

So far we have dealt entirely in rectangular coordinates in the λ and Z planes. Equation (46) however involves an inequality between λ and $Z(\lambda)$ in polar coordinates. This is essentially a question of conformality, and like most questions of this nature can be made to depend on a fundamental lemma of Schwarz.[4] For our purposes the following generalization of this lemma is convenient.

THEOREM DUE TO PICK (13): If the function W of γ has no essential singularities for values of γ within the circle K_γ, and takes on values which lie only in the interior of another circle K_W, then all non-euclidean distances, elements of arcs and arcs are shortened in the conformal mapping by $W(\gamma)$. If one such measure remains unchanged, all remain unchanged and W is a linear function of γ.

In this theorem a non-euclidean distance between two points γ_1, γ_2 is defined by the logarithm of a cross-ratio in the usual manner of projective measurement as follows:—Through the points γ_1, γ_2 construct a circle orthogonal to the circle K_γ and cutting it in the points h and k; the non-euclidean distance is then

$$\mathcal{P}(\gamma_1, \gamma_2) = \log \frac{\gamma_1 - h}{\gamma_1 - k} \cdot \frac{\gamma_2 - k}{\gamma_2 - h}. \tag{8}$$

This theorem of Pick also holds for the limiting case when the interior of the circle K_γ or K_W becomes a half-plane.

[4] H. A. Schwarz 1869. See also Bibliogr (13) and (14) Vol. II, p. 114.

Now W will be a "positive" function of γ if K_γ and K_W become the right half of the γ and W planes respectively. The further restriction that W be real for real values of γ makes the correspondence between W and γ on the one hand and a "positive real" function $Z(\lambda)$ and λ on the other hand complete.

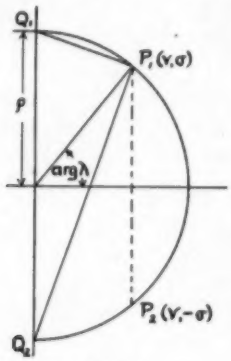


FIG. 5

All that remains is to interpret the non-euclidean distances. We shall do this only for the more interesting cases:—

(i) Two conjugate points λ, $\bar{\lambda}$ will transform into two other conjugate points Z, $\bar{Z}$ because of the condition of "reality." In Fig. 5 let $P_1(\lambda)$, $P_2(\bar{\lambda})$ be a pair of conjugate points and let

$$\lambda = \nu + i\sigma, \quad |\lambda| = \rho.$$

The circle through λ, $\bar{\lambda}$ orthogonal to the boundary of the right half plane will cut this boundary in the points $Q_1(0 + i\rho)$ and $Q_2(0 - i\rho)$

$$\begin{aligned} \mathfrak{P}(\lambda, \bar{\lambda}) &= \log \frac{(\nu + i\sigma - \rho)}{(\nu + i\sigma + \rho)} \cdot \frac{(\nu - i\sigma - \rho)}{(\nu - i\sigma + \rho)} \\ &= \log \frac{\nu^2 + (\sigma - \rho)^2}{\nu^2 + (\sigma + \rho)^2} \\ &= 2 \log \frac{P_1Q_1}{P_1Q_2}. \end{aligned} \tag{9}$$

As $\mathcal{P}(\lambda, \bar{\lambda})$; decreases towards zero (in absolute magnitude since sign does not enter into the idea of distance) the ratio P_1Q_1/P_1Q_2 must approach unity. A ratio of lengths will involve only angles and in Fig. 5 an approach of the ratio P_1Q_1/P_1Q_2 to unity is readily seen to mean a decrease in the angle P_1OP_2 towards zero, i.e. $|\arg \lambda|$ decreases.

(ii) Consider two points $\lambda_1\lambda_2$ on the real axis. The orthogonal circle in this case is the real axis itself. The non-euclidean distance, obtained by considering a limit, becomes

$$\mathcal{P}(\lambda_1, \lambda_2) = \log \frac{\lambda_1}{\lambda_2}. \tag{10}$$

From this we see that if for one value of λ_0 on the positive real axis we have $\lambda_1 = Z(\lambda_1) = Z_1$, this will be the only point for which such a relation holds unless $\lambda \equiv Z(\lambda)$. For if there be another point where $\lambda_2 = Z(\lambda_2) = Z_2$ we shall have

$$\mathcal{P}(\lambda_1, \lambda_2) = \log \frac{\lambda_1}{\lambda_2} = \log \frac{Z_1}{Z_2} = \mathcal{P}(Z_1Z_2),$$

which by the theorem of Pick is possible only if this is true for all values of λ. The same is obviously true if we substitute $L\lambda$ for λ (L being a real positive constant).

Hence we have

THEOREM VI. (i) if $Z(\lambda)$ is a "positive real" function $|\arg Z(\lambda)| \leq |\arg \lambda|$ for all values of λ satisfying $0 < |\arg \lambda| \leq \pi/2$. The equality signs can only hold simultaneously, unless they hold identically.

(ii) On the positive real axis (where $\arg \lambda = 0$) the equation $L\lambda = Z(\lambda)$ cannot be satisfied at more than one internal point. (L a real positive constant.)

THEOREM VI, COROLL. 1. The function $(Z(\lambda) - L\lambda)$, (which is not necessarily a positive function) cannot have more than one zero within the right-half plane.

Theorem VI may be interpreted in the light of the physical interpretation given to Theorem I. Thus if λ ber eplaced by a

" real positive " function of λ the resulting network will have a power factor closer to unity than the original network, for corresponding frequencies.

The converse of Theorem VI is at once evident for if $|\arg \lambda| \leq \pi/2$, λ lies in the right-half plane and $|\arg Z(\lambda)|$ being $\leq |\arg \lambda|$, $Z(\lambda)$ will likewise lie in the right-half plane. Furthermore when $\arg \lambda = 0$, $\arg Z(\lambda) = 0$, i.e. Z is a " real " function. Hence we have

THEOREM VII. If $|\arg Z(\lambda)| \leq |\arg \lambda|$ when $|\arg \lambda| \leq \pi/2$, $Z(\lambda)$ will be a " positive real " function.

It is now apparent that if $Z(\lambda)$ is a " positive real " function it will satisfy all the necessary conditions mentioned in Part I, section 3; for conditions $2(a)$ and $2(b)$ follow from Theorem II; condition (3) is merely $(4a)$ when $\mathrm{Re}\,\lambda = 0$; and condition $(4b)$ is fulfilled by Theorem VI.

Theorem V gives two sets of conditions equivalent to $(4a)$ which are easier to handle numerically and more readily interpreted electrically.

From Theorem VII we see that $(4b)$ is an alternative way of defining a " positive real " function.

We next proceed to show that a " positive real " function not only satisfies all the necessary conditions immediately derivable from physical considerations, but that these conditions are also *sufficient* in order that a network representation of the function may be found. For this purpose the above theorems will prove particularly useful.

PART III. SYNTHESIS OF NETWORKS WITH PRESCRIBED IMPEDANCE FUNCTION

1. *Functions with not more than Two Poles (or Zeros) in the Interior of the Left Half Plane*

With the more detailed formulation of the properties of " positive real " functions given in Part II we proceed to consider some of the simpler cases of such functions and their net-

work representations. These have been treated in part by Foster, Cauer and Fry, but the greater power of the methods now at our disposal will be abundantly clear even for these cases.

For convenience in calculation we use the conditions for a " positive real " function as stated in Theorem V. We may restrict ourselves to functions which have no poles or zeros on the axis of imaginaries for if such poles or zeros are present they may be removed in the manner demonstrated in Theorem IV; they correspond to pure reactance elements in series or parallel with each other and with the remaining network.

(*a*) *Functions with One Pole (or Zero)*

Consider the function [5]

$$Z(\lambda) = \frac{a_0 + a_1\lambda}{b_0 + b_1\lambda}, \tag{11}$$

$$\text{Re } Z(i\omega) = \frac{a_0b_0 + a_1b_1\omega^2}{b_0^2 + b_1^2\omega^2}. \tag{12}$$

We may arbitrarily assume a_0 positive without imposing any restrictions on $Z(\lambda)$. For Re$\cdot Z(i\omega) \geq 0$ we must have $b_0 \geq 0$ and $a_1b_1 \geq 0$. The condition for the pole requires $b_1 \geq 0$ hence also $a_1 \geq 0$.

Let us now plot Re $Z(i\omega)$ as a function of ω^2 (negative values of ω^2 are included in the figure for the sake of completeness although they do not enter into the present discussion). The curve is a hyperbola which takes one of two forms depending on whether $\frac{a_0}{b_0} <$ or $> \frac{a_1}{b_1}$ (see Fig. 6). In the case $\frac{a_0}{b_0} < \frac{a_1}{b_1}$ (Fig. 6*a*) the smallest value of Re $Z(i\omega)$ in the right half plane occurs at the origin. This value may be subtracted from $Z(\lambda)$ and still leave a " positive " function (Theorem III, Coroll. 1). Then

$$Z(\lambda) = \frac{a_0}{b_0} + \frac{(a_1b_0 - b_1a_0)\lambda}{b_0^2 + b_0b_1\lambda}$$
$$= R_1 + Z'(\lambda). \tag{13}$$

[5] This function was discussed by Fry, Bibliogr (4).

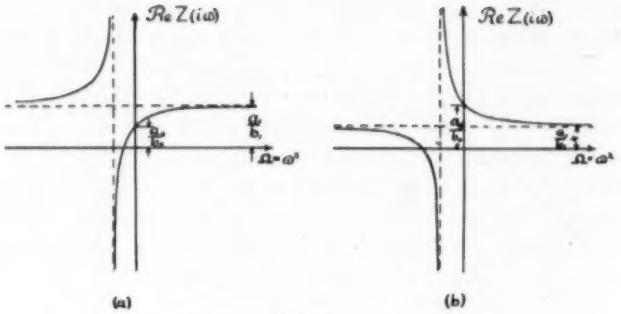


FIG. 6

$Z'(\lambda)$ may be immediately recognized as the impedance function of a network, or we may proceed by noting that it has a zero at the origin. Hence applying Theorem IV

$$\frac{1}{Z'(\lambda)} = \frac{b_0^2}{(a_1b_0 - b_1a_0)\lambda} + \frac{b_0b_1}{(a_1b_0 - b_1a_0)}. \tag{14}$$

The corresponding network is shown in Fig. 7 (*a*).

An exactly similar procedure is possible when $\frac{a_0}{b_0} > \frac{a_1}{b_1}$ (see Fig. 6*b*). Then

$$Z(\lambda) = \frac{a_1}{b_1} + \frac{(a_0b_1 - a_1b_0)}{b_0b_1 + b_1^2\lambda}$$

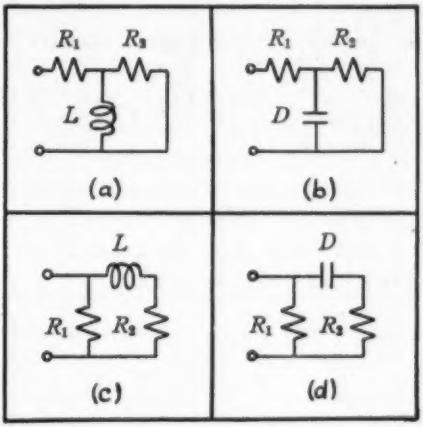
and so on. (The results are obtainable from those of section 40 by the substitution $\lambda' = 1/\lambda$ or by an interchange of the subscripts 0 and 1.) The corresponding network is shown in Fig 7 (*b*).

The network representation of (10) thus contains resistance and inductance or resistance and capacitance according as

$$\frac{a_0}{b_0} < \quad \text{or} \quad > \frac{a_1}{b_1} \tag{15}$$

When $\frac{a_0}{b_0} = \frac{a_1}{b_1}$, $Z(\lambda)$ obviously reduces to $R = \frac{a_0}{b_0} = \frac{a_1}{b_1}$.

A corresponding discussion which is dual to the above in the sense of Theorem 1, Coroll. 1 can be carried through with $1/Z(\lambda)$ giving rise to the networks in Fig. 7 (*c*) and (*d*). The networks



$$(a)\begin{cases} R_1 = \dfrac{a_0}{b_0} \\ R_2 = \dfrac{b_0 b_1}{(a_1 b_0 - a_0 b_1)} \\ L = \dfrac{b_0^2}{(a_1 b_0 - a_0 b_1)} \end{cases} \qquad (b)\begin{cases} R_1 = \dfrac{a_1}{b_1} \\ R_2 = \dfrac{b_0 b_1}{(a_0 b_1 - a_1 b_0)} \\ D = \dfrac{b_1^2}{(a_0 b_1 - a_1 b_0)} \end{cases}$$

$$(c)\begin{cases} R_1 = \dfrac{a_1}{b_1} \\ R_2 = \dfrac{a_0 a_1}{(a_1 b_0 - a_0 b_1)} \\ L = \dfrac{a_1^2}{(a_1 b_0 - a_0 b_1)} \end{cases} \qquad (d)\begin{cases} R_1 = \dfrac{a_0}{b_0} \\ R_2 = \dfrac{a_0 a_1}{(a_0 b_1 - a_1 b_0)} \\ D = \dfrac{a_0^2}{(a_0 b_1 - a_1 b_0)} \end{cases}$$

FIG. 7

of Figs. 7 (*c*) and (*d*) are *equivalent* to the networks of Figs. 7 (*a*) and (*b*) respectively.

It is interesting to note that the complete set of equivalent networks (having only two meshes) is obtained by combining the processes leading to Figs. 7*a* and 7*c* in different ways. For example if $\dfrac{a_0}{b_0} < \dfrac{a_1}{b_1}$ any value less than a_0/b_0 can be separated as a

series resistance (Fig. 7*a*) and thereafter the remaining network constructed as in Fig. 7 (*c*).

The above discussion disposes immediately of such forms as

$$\left.\begin{array}{l|l} \text{(i)}\ Z(\lambda) = \dfrac{a_0 + a_1\lambda + a_2\lambda^2}{b_0 + b_1\lambda}, & \text{(iii)}\ Z(\lambda) = \dfrac{a_0 + a_1\lambda + a_2\lambda^2}{b_1\lambda + b_2\lambda^2}, \\ \text{(ii)}\ Z(\lambda) = \dfrac{a_0 + a_1\lambda}{b_0 + b_1\lambda + b_2\lambda^2}, & \text{(iv)}\ Z(\lambda) = \dfrac{a_1\lambda + a_2\lambda^2}{b_0 + b_1\lambda + b_2\lambda^2}. \end{array}\right\} \quad (16)$$

The corresponding networks are shown in Fig. 8 for future reference. The coefficients in these functions may not be arbitrary positive constants as in (10), the additional conditions

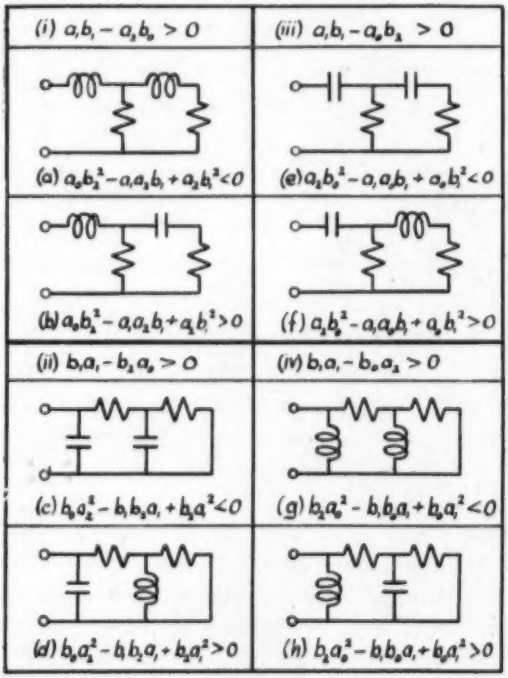


FIG. 8

which they must fulfill being shown in the Figure for each case. These conditions are readily obtained by applying the conditions for (10) to the reduced functions which result on separating the pole (or zero) on the imaginary axis from the forms (16).

It is interesting to note the dualities in Fig. 8. The forms (ii) and (iv) are derivable from (i) and (iii) respectively by the substitution $Z' = 1/Z$ while (iii) and (iv) are derivable from (i) and (ii) respectively by the substitution $\lambda' = 1/\lambda$ (compare Theorem I, Coroll. 1 and 2). A discussion of any one of them is therefore readily adaptable to all the others, by a mere change of notation.

Equivalents to the networks in Fig. 8 similar to those in Fig. 7 are also possible.

(b) Two Poles and Two Zeros not on the Boundary

Consider next the case

$$Z(\lambda) = \frac{a_0 + a_1\lambda + a_2\lambda^2}{b_0 + b_1\lambda + b_2\lambda^2}, \tag{17}$$

$$\text{Re } Z(i\omega) = \frac{a_0b_0 + (a_1b_1 - a_2b_0 - a_0b_2)\omega^2 + a_2b_2\omega^4}{b_0^2 + (b_1^2 - 2b_0b_2)\omega^2 + b_2^2\omega^4}. \tag{18}$$

The condition for the poles to have negative real part is given by

$$\frac{b_1}{b_0} \text{ and } \frac{b_2}{b_0} \geq 0. \tag{19}$$

In (18) the denominator cannot be negative for real values of ω (this is best seen by expressing it as the sum of squares). Hence the numerator must always be positive, i.e. it must have no real positive roots when considered as a function of ω^2. Hence we must have (when $b_0 > 0$) $a_0 > 0$, and either

$$a_2 > 0, \quad a_0b_2 - a_1b_1 + a_2b_0 \leq 0 \tag{20a}$$

or

$$(a_0b_2 - a_1b_1 + a_2b_0)^2 - 4a_0a_2b_0b_2 \leq 0. \tag{20b}$$

Assuming that the coefficients satisfy the conditions (19), and (20*a*)or (20*b*), let us consider the possible manners of variation

of Re $Z(i\omega)$ according to equation (18). Let us write

$$\text{Re}\, Z(i\omega) = R(\Omega) = \frac{A_0 + A_1\Omega + A_2\Omega^2}{B_0 + B_1\Omega + B_2\Omega^2}, \qquad (18a)$$

where $\Omega = \omega^2$.

The discussion of such functions is an elementary problem in algebra, and is usually carried through in the following way. From equation (18*a*)

$$(RB_0 - A_0) + (RB_1 - A_1)\Omega + (RB_2 - A_2)\Omega^2 = 0.$$

For real values of Ω we must have

$$(RB_1 - A_1)^2 - 4(RB_0 - A_0)(RB_2 - A_2) \geq 0,$$

i.e.

$$R^2(B_1^2 - 4B_0B_2) - 2R(A_1B_1 - 2A_0B_2 + A_2B_0) + (A_1^2 - 4A_0A_2) \geq 0. \qquad (21)$$

This means that R can or cannot (for real values of Ω) take on values lying between the roots of the expression (21) according as the sign of $(B_1^2 - 4B_0B_2)$ is negative or positive. (Note that this is identical with the sign of $b_1^2 - 4b_0b_2$, i.e. is determined by the poles of $Z(\lambda)$ being real or complex.)

Only in the case where the roots of this expression are equal or pure imaginary can R take on all values. Typical curves for R are shown in Fig. 9. In these figures the axes can be shifted parallel to themselves and reverse direction in any manner provided only R remains positive for positive values of Ω. Thus for example Figs. 9 (*b*) and (*d*) are of the same type, but with the direction of the R axis reversed.

As in the discussion of $Z(\lambda)$ defined by (11) the procedure depends on the frequency at which the least value of the 'resistance' $R(\Omega)$ occurs. For this purpose we need consider only positive values of Ω in Fig. 9 since these alone correspond to points on the boundary of the right-half plane (or to real frequencies).

If this least value occurs for $\Omega = 0$ or ∞, as for example in Figs. 9 (a) and (b) the procedure is essentially the same as before. To illustrate, let the minimum value of $R(\Omega)$ occur for $\Omega = 0$

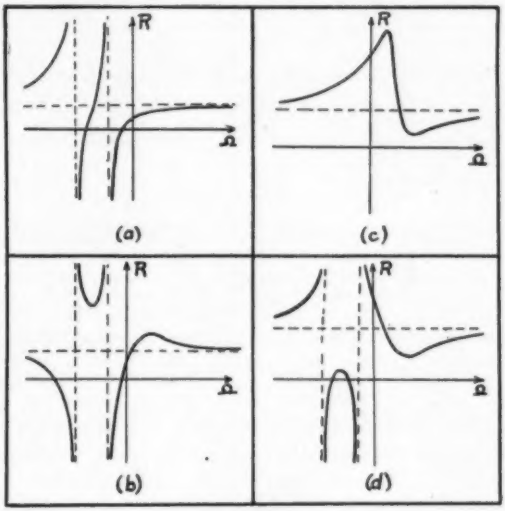


FIG. 9

$$\text{i.e.}\ \frac{a_0}{b_0} \leq R(\Omega) \quad \text{for} \quad \Omega > 0. \tag{22}$$

Then

$$\begin{aligned} Z(\lambda) &= \frac{a_0 + a_1\lambda + a_2\lambda^2}{b_0 + b_1\lambda + b_2\lambda^2} \\ &= \frac{a_0}{b_0} + \frac{(a_1b_1 - a_0b_1)\lambda + (a_2b_0 - a_0b_2)\lambda^2}{b_0(b_0 + b_1\lambda + b_2\lambda^2)} \\ &= \frac{a_0}{b_0} + Z'(\lambda), \end{aligned}$$

in which the function $Z'(\lambda)$ is a " positive " function under the assumption (22) by Theorem III, Coroll. 1. Furthermore this

form of $Z'(\lambda)$ has already been discussed (Fig. 8 (iv)). Consequently $Z(\lambda)$ can be represented by such a network together with a series resistance, as shown in Fig. 10 A and B.

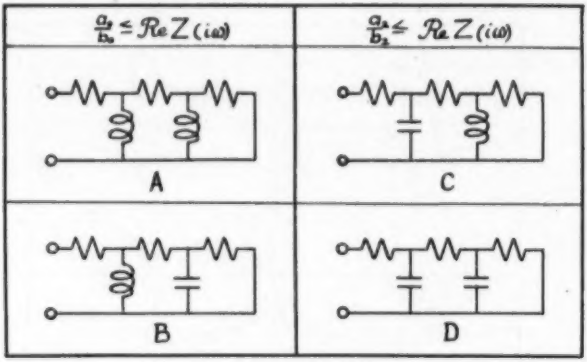


FIG. 10

The case when $R(\omega)_{\min}$ occurs for $\omega = \infty$ needs no further discussion, the corresponding networks being shown in Fig. 10 C and D.

There remains the case when $R(\omega)_{\min}$ occurs for some finite value $\omega = \sigma$. Proceeding by analogy, let this value R_σ be separated out. It will be a solution of (21) (with the equality sign). Let us write

$$Z(\lambda) = R_\sigma + \frac{a_0' + a_1'\lambda + a_2'\lambda^2}{b_0 + b_1\lambda + b_2\lambda^2} = R_\sigma + Z'(\lambda). \qquad (23)$$

The function $Z'(\lambda)$ now does not have a zero on the axes of imaginaries, as in equation 28*a*, but will be characterized by the fact that for some value of frequency ($\lambda = i\sigma$) $Z'(\lambda)$ will be pure imaginary. Physically this means that the corresponding network behaves like a pure reactance at this frequency. This can only mean that that part of the network which contains resistance is short circuited by (i.e. is in parallel with) a reactive branch which is in resonance for this fre-

quency; in series with this is a pure reactance which may be either positive or negative, since the pure imaginary part may have either sign. The only type of element which can meet this requirement is inductance with mutual coupling. Accordingly let us consider the network of Fig. 11 as a possible repre-

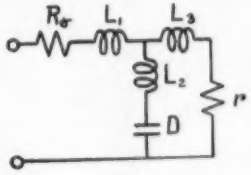


FIG. 11

sentation of the function (17) under these circumstances. In this network L_1, L_2, L_3, are an equivalent representation of two inductances with mutual coupling: one of them may be negative, with the necessary and sufficient condition for physical realizability

$$L_1 + L_2 \geq 0; \quad L_1L_2 + L_2L_3 + L_3L_1 \geq 0 \tag{24}$$

whence it follows that $L_2 + L_3 > 0$ and $L_3 + L_1 > 0$.

Now considering the network remaining after R_σ has been removed from the network of Fig. 11, its impedance function is readily found to be

$$Z''(\lambda) = \frac{(L_1L_2 + L_2L_3 + L_3L_1)\lambda^3 + (L_1 + L_2)r\lambda^2 + (L_1 + L_3)D\lambda + rD}{(L_2 + L_3)\lambda^2 + r\lambda + D}. \tag{25}$$

Comparison with $Z'(\lambda)$ in equation (23) makes it apparent that

$$L_1L_2 + L_2L_3 + L_3L_1 = 0 \tag{26}$$

which is the condition that the two coils of which L_1, L_2, L_3 are the equivalent shall be perfectly coupled.

Further relations which have to be satisfied are

$$r = \frac{a_0'}{b_0};\quad D = \frac{a_0'}{b_1}\left(= b_0 \times \frac{a_0'}{b_1 b_0}\right)$$

$$\left.\begin{aligned} L_1 + L_2 &= \frac{a_2'}{b_1} \\ L_2 + L_3 &= b_2 \cdot \frac{a_0'}{b_1 b_0} \\ L_1 + L_3 &= \frac{a_1'}{b_0} \end{aligned}\right\} \text{i.e.} \left\{\begin{aligned} L_1 &= \frac{-a_0'b_2 + a_1'b_1 + a_2'b_0}{2b_1b_0} \\ L_2 &= \frac{a_0'b_2 - a_1'b_1 + a_2'b_0}{2b_1b_0} \\ L_3 &= \frac{a_0'b_2 + a_1'b_1 - a_2'b_0}{2b_1b_0} \end{aligned}\right\} \quad (27)$$

Consequently the relation (26) becomes

$$\begin{aligned} 0 &= L_1L_2 + L_3(L_1 + L_2) \\ &= (-a_0'b_2 + a_1'b_1 + a_2'b_0)(a_0'b_2 - a_1'b_1 + a_2'b_0) \\ &\qquad + 2a_2'b_0(a_0'b_2 + a_1'b_1 - a_2'b_0) \\ &= 4a_0'a_2b_0b_2 - (a_0'b_2 - a_1'b_1 + a_2'b_0)^2. \end{aligned} \quad (28)$$

Comparing this with equation (18) and remembering that $\operatorname{Re} Z'(i\omega)$ has a double zero when $\omega^2 = \sigma^2$, it is evident that condition (28) is satisfied. Hence $Z(\lambda)$ may be represented by the network in Fig. 11 by going through the steps indicated in equations (23) and (27).

We have thus established a procedure whereby a network representation may be found for any " real positive " function of the form given in equation (17). The network will have one of the forms shown in Figs. 10 and 11.

2. *General Procedure*

The methods exemplified in the last section are, with very little modification, applicable to the general case of a " positive real " function with a finite number of poles and zeros. The general procedure for finding a network representation of such a function will be a step by step process; at each step the function is decomposed into a part which can immediately be represented by an element of the network, and another part which is again a positive real function. The new function should be simpler than the previous function if any progress is to be made towards

complete representation; this will show itself in a decrease in the number of poles and zeros. (The number of poles being always equal to the number of zeros.)

Let the given function be $Z(\lambda)$. Poles and zeros on the axis of imaginaries can immediately be removed by Theorem IV and corresponding network elements constructed. When this has been repeated often enough we are left with a " positive real " function which is either a constant or has all its poles and zeros in the interior of the left-half plane. Let this function be $Z_1(\lambda)$.

Next examine the real part of $Z_1(\lambda)$ on the axis of imaginaries.

We remark in passing the Re $Z(i\omega)$ is an even function of ω while Im $Z(i\omega)$ is odd, in other words we may always write

$$Z(i\omega) = R(\omega^2) + i\omega N(\omega^2). \tag{29}$$

Since Re $Z(i\omega)$ cannot be negative it must have a lower limit ≥ 0. Let this minimum value be R_1 and occur when $\omega = \sigma$ (σ real). Then clearly $Z_2(\lambda) = Z_1(\lambda) - R_1$ will have its real part ≥ 0 on the boundary and regular in the right-half plane, i.e. $Z_2(\lambda)$ will be a " positive real " function. Moreover $Z_2(i\sigma)$ will be pure imaginary. Let

$$Z_2(i\sigma) = iX. \tag{30}$$

Now if $\sigma = 0$ or ∞, $X = -X = 0$; hence we can immediately conclude that $Z_2(\lambda)$ has a zero at $\lambda = 0$ or ∞ respectively. We could therefore write

$$\frac{1}{Z_2(\lambda)} = \frac{D}{\lambda} + \frac{1}{Z_3(\lambda)} \tag{31a}$$

or

$$\frac{1}{Z_2(\lambda)} = L\lambda + \frac{1}{Z_3(\lambda)} \tag{31b}$$

respectively; in either case $Z_3(\lambda)$ will have one less zero and pole than $Z_2(\lambda)$ (or $Z_1(\lambda)$).

Consider next the case when σ is not zero or infinite.

We proceed to make the points $\lambda = \pm i\sigma$ zeros of a new function simply derivable from $Z_2(\lambda)$ by a step corresponding to

the realization of an element in the network representation. Note that if

$$Z_2(i\sigma) = iX,$$

then

$$\frac{1}{Z_2(i\sigma)} = -i\frac{1}{X}.$$

Two conditions may arise; $X/\sigma = N(\sigma^2)$ (equation (29)) may be either positive or negative. We consider these two possibilities separately.

(i) Let $X/\sigma = N(\sigma^2) = L_1$ be negative.

Then we can write

$$Z_2(\lambda) = L_1\lambda + W(\lambda), \tag{32}$$

where $W(\lambda)$ is a "positive real" function and has zeros at $\lambda = \pm i\sigma$. For $-L_1\lambda$ is a "positive real" function and $W(\lambda) = Z_2(\lambda) - L_1\lambda$ is thus the sum of two "positive" functions; also $W(i\sigma) = Z_2(i\sigma) - iX = 0$. (Note: $W(\lambda)$ has a pole at infinity.) Hence we may write

$$\frac{1}{W(\lambda)} = \frac{K\lambda}{\lambda^2 + \sigma^2} + \frac{1}{W'(\lambda)}, \tag{33}$$

where $W'(\lambda)$ is a "positive real" function which, like $W(\lambda)$, has a pole at infinity.

Let

$$W'(\lambda) = L_3\lambda + Z_3(\lambda). \tag{34}$$

If we further write

$$L_2 = \frac{1}{K}, \quad D = \frac{\sigma^2}{K} \tag{35}$$

the relation between $Z_2(\lambda)$ and $Z_3(\lambda)$ in terms of a network representation is shown in Fig. 12. Comparing this with Fig. 11 and formula (25) and remembering that $Z_2(\lambda)$ has no pole at $\lambda = \infty$ we have as the algebraic relation

$$Z_2(\lambda) = \frac{(L_1 + L_2)\lambda^2 Z_3(\lambda) + (L_3 + L_1)D\lambda + DZ_3(\lambda)}{(L_2 + L_3)\lambda^2 + Z_3(\lambda)\cdot\lambda + D}. \tag{36}$$

Also L_2, L_3 are positive by (33), (34) and (35) (cf. equation 24) hence the network is realizable.

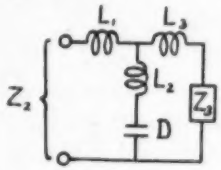


FIG. 12

(ii) Let $X/\sigma = N(\sigma^2)$ be positive. Then we may consider

$$Y_2(\lambda) = \frac{1}{Z_2(\lambda)} \quad \text{where} \quad Y_2(i\sigma) = -i\frac{1}{X} = -i\sigma\frac{1}{\sigma^2 N(\sigma^2)}$$

and

$$C_1 = -\frac{1}{\sigma^2 N(\sigma^2)}$$

is negative and proceed in an exactly analogous manner as with the preceding case. Translating the results into network elements, however, the network will be reciprocal to that in Fig. 12. This is shown in Fig. 13, L_1 being replaced by C_1, L_2 by C_2, D

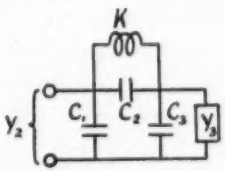


FIG. 13

by K and L_3 by C_3. The relation between $Y_2(\lambda)$ and $Y_3(\lambda)$ will be

$$Y_3(\lambda) = \frac{(C_1 + C_2)\lambda^2 Y_3(\lambda) + (C_3 + C_1)K\lambda + KY_3(\lambda)}{(C_2 + C_3)\lambda^2 + Y_3(\lambda)\cdot\lambda + K}. \qquad (37)$$

The network elements in Fig. 13 are not directly realizable since they contain the negative capacity C_1. Equation (37)

may, however, be rewritten

$$Z_2(\lambda) = \frac{(C_3 + C_2)\lambda^2 Z_3'(\lambda) + \lambda + KZ_3'(\lambda)}{(C_1 + C_2)\lambda^2 + (C_3 + C_1)KZ_3'(\lambda)\cdot\lambda + K} \quad (37a)$$

where $Z_3'(\lambda) = 1/Y_3(\lambda)$

which is of exactly the same form as (36). Let us write down the necessary relations between the coefficients of (37*a*) and (36) in order that Z_3 and Z_3' may be identical. They are

$$D = \frac{1}{C_3 + C_1}, \quad (38a)$$

$$\left.\begin{aligned} L_1 + L_2 &= \frac{1}{K}\frac{C_2 + C_3}{C_3 + C_1}, \\ L_2 + L_3 &= \frac{1}{K}, \\ L_1 + L_3 &= \frac{1}{K}\frac{C_1 + C_2}{C_3 + C_1}, \end{aligned}\right\} \quad (38b)$$

we have also (analogous to 37*a* and 31*a*)

$$\left.\begin{aligned} \frac{K}{C_2} &= \sigma^2, \\ C_1C_2 + C_2C_3 + C_3C_1 &= 0, \end{aligned}\right\} \quad (38c)$$

and from 38*a*

$$\left.\begin{aligned} L_1 &= \frac{1}{K}\frac{C_3}{C_3 + C_1}, \\ L_2 &= \frac{1}{K}\frac{C_2}{C_3 + C_1}, \\ L_3 &= \frac{1}{K}\frac{C_1}{C_3 + C_1}, \end{aligned}\right\} \quad (38d)$$

with the relations from (38*c*)

$$\frac{D}{L_2} = \frac{K}{C_2} = \sigma^2;$$

$$L_1L_2 + L_2L_3 + L_3L_1 = \frac{C_3C_2 + C_2C_1 + C_1C_3}{K^2(C_3 + C_1)^2} = 0. \quad (38e)$$

It follows that the fictitious network of Fig. 13 can be made equivalent to the physically realizable network of Fig. 12 by means of the equations (38*a*) and (38*d*). Note that the signs of L_1, L_3 and C_1, C_3 are interchanged. Hence in case (ii) L_1 and L_2 are positive, and L_3 is negative.

Furthermore, since we now know the structure of the network it is evident that $Z_2(i\sigma) - iL_1\sigma$ must be zero. Consequently $L_1 = N(\sigma^2)$, and the procedure of Case (i) may be followed irrespective of the sign of $N(\sigma^2)$. (At one stage a non-" positive " function will be encountered but the final $Z_3(\lambda)$ will be a " real positive " function.)

Examining the reduction of poles and zeros by this process, we note that the first step (equation 32) increases the number by 1. The next step (equation 33) decreases the number by 2, while the last step (equation 34) effects a further reduction of 1. In all, thus, if R_1 occurs for $\lambda = i\sigma \neq 0$ or ∞, $Z_3(\lambda)$ will have two poles and two zeros less than $Z_1(\lambda)$.

This procedure can now be applied to all cases and repeated until the final function has no poles and no zeros, i.e. is a positive constant corresponding to a pure resistance. It leads uniquely to a network representation of the original " positive real " function.

The calculation of the network is made entirely in terms of the values of successive impedance functions on the axis of imaginaries. Each reduction of the function is effected by the removal of a zero (or pole) of the function on the axis of imaginaries. If a function does not have such a zero on the imaginary axis it is shown always to be possible to derive from it a function which will have such a zero; the removal of this zero together with the related steps also corresponds to the calculation of realizable elements of the corresponding network.

For more ready reference we summarize the different steps with the corresponding network interpretations in Table I, and state the conclusion as

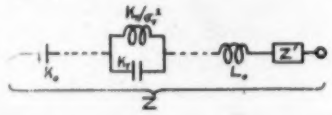
TABLE I

GENERAL PROCEDURE FOR CONSTRUCTING A NETWORK REPRESENTATION OF ANY GIVEN "POSITIVE" REAL FUNCTION

(*a*) Remove all poles on the axis of imaginaries:—

$$Z(\lambda) = L_0\lambda + \Sigma \frac{K_r\lambda}{\lambda^2+\sigma_r^2} + Z'(\lambda)$$

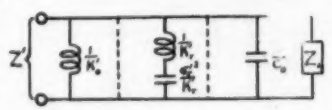
(Theor. IV)



(*b*) Remove all zeros on the axis of imaginaries:—

$$\frac{1}{Z'(\lambda)} = C_0\lambda + \Sigma \frac{K_r'\lambda}{\lambda^2+\sigma_r'^2} + \frac{1}{Z_1(\lambda)}$$

(Theor. IV, Cor. 1)



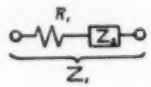
Repeat (*a*) and (*b*) until no poles or zeros of $Z_1(\lambda)$ lie on the axis of imaginaries.

(*c*) Examine $R(\omega^2)$ in

$$Z_1(i\omega) = R(\omega^2) + i\omega N(\omega^2)$$

Let R_1 be the least value of $R(\omega^2)$ for positive values of ω^2 and let

$$R_1 = R(\sigma^2)$$
$$Z_1(\lambda) = R_1 + Z_2(\lambda)$$



(*d*) Write $L_1 = N(\sigma^2)$ (σ found in (*c*))

$$Z_2(\lambda) = L_1\lambda + W(\lambda)$$

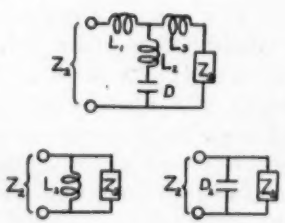
$W(\lambda)$ has a zero at $\lambda = \pm i\sigma$ and hence

$$\frac{1}{W(\lambda)} = \frac{\lambda}{L_2\lambda^2+D} + \frac{1}{L_3\lambda+Z_3(\lambda)}$$

As special cases when $\sigma = 0$ or ∞, this reduces to $L_1 = 0$, $L_3 = 0$

$$\frac{1}{Z_2(\lambda)} = \frac{1}{L_2\lambda} + \frac{1}{Z_3(\lambda)}$$

$$\frac{1}{Z_2(\lambda)} = \frac{D_3}{\lambda} + \frac{1}{Z_3(\lambda)}$$



THEOREM VIII. To every " positive real " function with a finite number of poles and zeros corresponds a finite physically realizable network. Hence the terms " *positive real* " *function with a finite number of poles and zeros* and *impedance function of a finite network* with lumped linear passive elements are synonymous.

3. *Discussion of Networks Obtained*

A discussion of the general type and properties of the network obtained by the process summarized in Table I, as well as of certain special cases, is of interest.

It will be noticed first of all that the network always has a ladder structure. The ladder structure is thus capable of realizing the most general form of driving point impedance.

In general mutual inductances will be involved but this mutual inductance if present occurs always between self-inductances in adjacent meshes; in certain cases mutual inductance may be entirely absent, viz. when procedure (c), Table I, involves only minimum values R_1 at $\lambda = 0$ or ∞ throughout.

The procedure is carried out in such a way that mutual inductances, when they occur, do so with a coefficient of coupling unity; this apparently involves an ideal physical condition not generally inherent in the given impedance function. If, however, the function $Z(\lambda)$ has a pole at $\lambda = \infty$ the necessity for perfect coupling is only apparent, since a pole at $\lambda = \infty$ corresponds to a self-inductance L_0 in series with the perfectly coupled inductance L_a in Fig. 14. But this separation of L_0 and L_a is

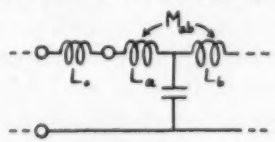


FIG. 14

quite arbitrary, the only limitation being that L_0 is the largest inductance which can be so separated and still leave the remaining network physically realizable. If L_0 (or any part of it) be

combined with L_a into an inductance L_a', coupled with the inductance L_b then obviously $L_a' \cdot L_b > M_{ab}^2$ and the coupling will be " loose." This condition of " loosening " will simultaneously extend throughout the remaining network. Only in such cases where the condition of " perfect " coupling is indeed inherent in the given impedance function because of the absence of a pole at $\lambda = \infty$ will this condition be forced on the corresponding network.

In fact a network structure like that indicated in Fig. 14 would under certain conditions be eminently suitable for practical construction. The series resistances are in the proper position to be combined with the inductances, while in the parallel branches only capacities, whose series resistance is negligible, occur.

The whole question of practical usability, however, depends on the special conditions to be met in the particular application. Its full discussion would of necessity involve the consideration of equivalent networks.

In this connection it may also be pointed out that the dual circuit met with in Fig. 13 could sometimes be used, namely in such cases where the negative capacity occurs in parallel with a larger positive capacity (corresponding to a zero at $\lambda = \infty$) so that the combination will be positive.

At each stage in the general procedure described in the preceding section the determination of a given number of independent elements in the network is accompanied by a corresponding reduction in the number of independent constants in the function which remains. The resulting network therefore does not contain any syperfluous elements. Let us consider this property of the network in greater detail.

By an independent network element is meant one whose value does not depend on the values of other elements; thus in Fig. 12 the elements of the group $L_1L_2L_3$ are not independent since their values are connected by the relation

$$L_1L_2 + L_2L_3 + L_1L_3 = 0 \qquad (26)$$

Similarly by the number of independent constants in the function is meant the number of such constants necessary to determine the function.

A function is completely determined by prescribing its poles, its zeros, and an arbitrary factor. Poles and zeros in the interior of the left-half plane are either real or occur in conjugate complex pairs. Hence to each such pole or zero in the interior of the left-half plane corresponds one independent constant. A pair of poles (or zeros) on the imaginary axis involves only one constant, while poles or zeros at the origin or infinity introduce no new constants. The number of independent constants in the function is thus

$$\mu = (n + m + h + 1) \tag{39}$$

where

n = number of poles in the interior of left-half plane,

m = number of zeros in the interior of left-half plane,

h = number of *pairs* of poles and zeros on the imaginary axis (excluding 0 and ∞).

We have stated this result in this way because a mere count of the number of coefficients in numerator and denominator is not always sufficient. To every conjugate pair of poles or zeros on the imaginary axis corresponds in general a relation among the coefficients. In the extreme case where all zeros (or poles) lie on the imaginary axis, these relations are simply expressed by stating that the coefficients of either the even or odd powers of λ in the numerator (or denominator) are zero. In intermediate cases the relations are not so easily seen and therefore likely to be missed unless a detailed examination of the poles and zeros is made.

With these facts in mind it is easy to trace the reduction of the independent constants in the function brought about by the removal of the various groups of elements in Table I from the network.

In special cases additional relations besides those mentioned

above may exist among the coefficients, making possible a still further reduction in the number of elements in the network. The possibility of making this reduction will not always be evident from the general procedure of Table I. A complete discussion of such possibilities would involve a discussion of all equivalent networks which will not be attempted here.

Fosters "Reactance Theorem" is readily seen to be only a particular case of the general procedure in which all the poles and zeros lie on the axis of imaginaries. For the other cases of networks having only two kinds of elements the procedure of Table I leads to the types of structure shown in Fig. 1 (*b*) and (*c*) in which either inductance coils or condensers are replaced by resistances.

Part IV—Extension of Method to the Determination of Certain Equivalent Networks

1. *Change in the Order of Procedure*

Very slight consideration will show that the order of the steps (*a*), (*b*), (*c*) and (*d*) specified in Table I, is not essential. Thus, for example, in the simple case where all the poles and zeros lie on the axis of imaginaries any pole or any zero may be removed first and thereafter again any pole or any zero. The networks shown in Fig. 1 are special cases when a definite order is adopted. Fig 1 (*a*) results when all poles of $Z(\lambda)$ are removed in succession, i.e. the network given by Table I; (*b*) when all zeros are removed in succession; (*c*) results when we remove first the pole at infinity, then the zero at infinity and so on alternately; and (*d*) when poles and zeros are removed alternately at the origin. The order may, however, be entirely haphazard. In fact only "part of a pole" i.e. a term representing a pole at the same point but with residue which is a positive proper fraction of the corresponding residue of $Z(\lambda)$, may be separated, the remaining positive function can thereafter be represented in any of the ways shown in Fig. 1 (*b*), (*c*) or (*d*); this would give a network with superfluous elements.

When we are dealing with a more general type of function, the order of removing poles and zeros on the imaginary axis is still arbitrary until the procedure of Table I (*c*) becomes necessary. Obviously if there are any zeros on the axis of imaginaries these will also be points at which Re $Z(i\omega)$ is zero; consequently step (*d*) becomes equivalent to (*b*) until all zeros on the axis of imaginaries have been removed. The presence of a pole in $Z_1(\lambda)$ on the imaginary axis however will not alter the value of R_1 in step (*c*). In the next step (*d*) therefore $Z_2(\lambda)$ will be a "positive" function and nothing will be altered in the argument concerning the validity of the procedure (*d*). In the event of a pole at infinity having been left in $Z_2(\lambda)$ we shall obtain $L_1L_2 + L_2L_3 + L_3L_1 > 0$ instead of $= 0$, which obviously still fulfills the physical requirements. In fact in this case L_1 is exactly equal to $L_0 + L_1$ (Table I (*a*) and (*d*)). Poles on the imaginary axis can therefore be left in the function or separated at will at any stage in the procedure. Another alternative is to treat $1/Z(\lambda)$ as an admittance; step (*c*) will correspond to the separation of a parallel conductance from $Z(\lambda)$; in this case zeros of $Z(\lambda)$ do not necessarily have to be first removed, but poles (i.e. zeros of $1/Z(\lambda)$) must.

2. *Consideration of "imaginary frequencies"*

In considering a network representation of "positive real" functions with two poles and two zeros. we were lead by certain physical considerations to the network of Fig. 12, which in turn lead to equations (27) and the relation (28). Now, while our physical reasoning depended on the fact that $R(\Omega)$ (equation 18*a*) had its minimum value for a positive value of Ω, equations (27) have no reference to this value of Ω. The necessary relation (28) will hold whenever $R'(\Omega)$ has a double zero, and in order that equations (27) may lead to realizable elements, it is only necessary that $Z'(\lambda)$ (equation (23)) shall be a "positive real" function. Both these conditions are satisfied if a stationary value of $R(\Omega)$ equal to R_r (say) be deducted from $Z(\lambda)$, and if

$$R_r \leq R(\Omega) \text{ for positive values of } \Omega. \tag{40}$$

This therefore gives us an alternative procedure in certain cases leading to equivalent networks. E.g. if $R(\Omega)$ varies in the manner shown in Fig. 10 (*d*) R_ν may be taken as the maximum occuring between the two vertical asymptotes in this Figure. In this case L_2 in Fig. 11 would be negative, and physical conceptions of 'resonance' would need to be extended to give a physical interpretation to this procedure. The type of variation of $R(\Omega)$ shown in Fig. 10 (*a*) on the otherhand will never permit this kind of procedure; it corresponds to networks with two kinds of elements only.

It will now be shown that a similar procedure can be justified in the general case.

Negative values of Ω correspond to imaginary values of ω which in turn means real values of λ. If therefore we write

$$Z(\lambda) = R(-\lambda^2) + \lambda N(-\lambda^2) \tag{41}$$

the interpretation in terms of real values of λ can be carried over from the discussion for pure imaginary values without any difficulty.

Let a stationary value of $R(-\lambda^2)$ satisfying (40) occur when $\lambda^2 = \nu^2$, and let $R_\nu = R(-\nu^2)$. Then if

$$Z_1(\lambda) = R_\nu + Z_2(\lambda) \tag{42}$$

it is clear that $Z_2(\lambda)$ will be a "positive real" function in virtue of (48). Now if

$$Z_2(\lambda) = (R(-\lambda^2) + R_\nu) + \lambda N(-\lambda^2)$$

we know that

$$Z_2(\nu) = \nu N(-\nu^2).$$

Let us write

$$L_1 = N(-\nu^2)$$

and

$$Z_2(\lambda) = L_1\lambda + W(\lambda). \tag{43}$$

Then it follows from equation (41) and (42) that

$$W(\lambda) = [R(-\lambda^2) - R_\nu] + \lambda[N(-\lambda)^2 - L_1], \tag{43a}$$

i.e.

$$W(\nu) = 0 \quad \text{and} \quad L_1 = Z_2(\nu)\nu. \tag{43b}$$

It is further clear that equations (43*b*) holds for $\lambda = \pm \nu$. Hence W will *not* be a "positive" function, since it has the zero $\lambda = \nu$ in the right-half plane. By Theorem VI, Coroll. 1 however we know that this is the only zero of $W(\lambda)$ within the right-half plane. Note further that

$$\text{Re } W(i\omega) \geq 0 \tag{44}$$

in virtue of (40).

If therefore we separate the zero's at $\lambda = \pm \nu$ from $W(\lambda)$ by a partial fraction expansion of $1/W(\lambda)$, the remaining function may be a positive real function.

For this purpose we examine the residues of the poles of $1/W(\lambda)$ at $\lambda = \pm \nu$. We may rewrite (43*a*)

$$W(\lambda) = R'(-\lambda^2) + \lambda N'(-\lambda^2)$$

which, in virtue of (43*b*) and the fact that R_ν is a stationary value of $R(-\lambda^2)$, can again be written

$$W(\lambda) = (\lambda^2 - \nu^2)^2 \cdot r(-\lambda^2) + \lambda(\lambda^2 - \nu^2) \cdot n(-\lambda^2)$$

Hence the residue at $\lambda = \nu$ is

$$\left[\frac{\lambda - \nu}{W(\lambda)}\right]_{\lambda=\nu} = \frac{1}{2\nu^2 n(-\nu^2)} = -\frac{K}{2} \text{ (say)}. \tag{45}$$

The residue at $\lambda = -\nu$ obviously has the same value.[6]

Consequently we may write

$$\frac{1}{W(\lambda)} = \frac{-\frac{K}{2}}{\lambda + \nu} + \frac{-\frac{K}{2}}{\lambda - \nu} + \frac{1}{W_1(\lambda)}$$

$$= \frac{K\lambda}{-\lambda^2 + \nu^2} + \frac{1}{W_1(\lambda)}. \tag{46}$$

Since $\text{Re } \dfrac{1}{W(i\omega)} \geq 0$ and the first term of the right-hand side

[6] This property is the exact analogy to the property that for a "positive real" function the residue at a pole on the imaginary axis is a positive real quantity.

of (46) is pure imaginary for $\lambda = i\omega$, it follows from (44) that $\operatorname{Re} \frac{1}{W_1(i\omega)} \geq 0$. But $W_1(\lambda)$ has no zeros in the right-half plane and no zeros on the imaginary axis. Consequently by Theorem V, Coroll. 1, $W_1(\lambda)$ is a " positive real " function. The process is completed by writing

$$W_1(\lambda) = L_3\lambda + Z_3(\lambda). \tag{47}$$

Also let

$$L_2 = -\frac{1}{K}, \quad D = \frac{\nu^2}{K}. \tag{48}$$

It is now only necessary to show that the reduction of $Z_2(\lambda)$ to $Z_3(\lambda)$ corresponds to the calculation of elements in a physically realizable circuit.

Referring to equation (43*b*), we have

$$L_1 = \frac{Z_2(\nu)}{\nu}, \tag{43b}$$

where, if ν is positive (and real), $Z_2(\nu)$ will be positive and real, since $Z_2(\lambda)$ is a " positive real " function. Consequently L_1 is positive.

Furthermore L_3, equation (47), is positive since Z_3 is a " positive real " function with a pole at infinity; and the relation

$$L_1L_2 + L_2L_3 + L_3L_1 = 0 \tag{26}$$

can be shown to hold in exactly the same manner as for Figs. 11 and 12, consequently L_2 is negative and D is positive. The process therefore corresponds to a determination of the elements L_1, L_2, L_3, D, of a realizable network connected in the manner of Fig. 12.

3. *Relation to Equivalent Networks Obtained by Affine Transformation of Quadratic Forms*

Equivalent networks have been derived from a given network by Howitt, (6) by the application of an affine transformation to the quadratic forms[7] connected with the network. Cauer (5) has also

[7] Corresponding to the stored electromagnetic and electrostatic energies and the rate of dissipation of energy in heat.

mentioned this method. Such a transformation can be accomplished by the matrix multiplication

$$C' \cdot A \cdot C = A', \tag{49}$$

where A is the matrix $\|a_{rs}\|$ of the system of equations

$$\left.\begin{aligned} E &= a_{11}I_1 + a_{12}I_2 + a_{13}I_3 + \cdots a_{1n}I_n \\ 0 &= a_{21}I_1 + a_{22}I_2 + a_{23}I_3 + \cdots a_{2n}I_n \\ &\cdot \qquad\quad \cdot \quad \cdot \quad \cdot \\ &\cdot \qquad\quad \cdot \quad \cdot \quad \cdot \\ 0 &= a_{n1}I_1 + a_{n2}I_2 + \cdots \qquad a_{nn}I_n \end{aligned}\right\} \tag{50}$$

for the given network; in these equations the I_s are a set of independent circulating mesh currents and

$$a_{rs} = L_{rs}\lambda + R_{rs} + D_{rs}\lambda^{-1} \tag{51}$$

are generalized self- $(r = s)$ and mutual- $(r \neq s)$ impedances of the sth mesh.

A' is a similar matrix $\|a_{rs}'\|$ for an equivalent network.

C is the transformation matrix $\|c_{rs}\|$ with c_{rs} real constants and

$$c_{11} = 1, \quad c_{1s} = 0 \quad (s \neq 1). \tag{52}$$

C' is the matrix $\|c_{sr}\|$ conjugate to C.

If we denote the determinant of A by D, the determinant of A' by D', and the determinant of C by Δ we have

$$D' = \Delta^2 D \tag{53}$$

and in virtue of (52) the first minor

$$\frac{\partial D'}{\partial a_{11}'} = \Delta^2 \frac{\partial D}{\partial a_{11}}. \tag{54}$$

It is on account of this property that the impedance function is invariant since

$$Z'(\lambda) = \frac{D'}{\dfrac{\partial D'}{\partial a_{11}'}} = \frac{D}{\dfrac{\partial D}{\partial a_{11}}} = Z(\lambda). \tag{55}$$

This property is however equally true of the determinants $|L_{rs}|$, $|R_{rs}|$, $|D_{rs}|$ and their first minors with respect to L_{11}, R_{11}, D_{11} respectively (equation (51)).

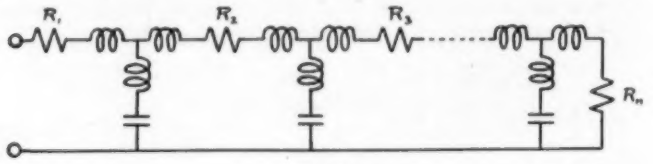


FIG. 15

Turning now to the networks obtained by the procedure outlined in the preceding sections we note that they always have the ladder structure indicated in Fig. 15 with no resistance in the parallel branches; consequently

$$|R_{rs}| = \begin{vmatrix} R_1 & 0 & 0 & \cdot & \cdot & 0 \\ 0 & R_2 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & R_3 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & R_n \end{vmatrix} \tag{56}$$

and the first minor

$$\frac{\partial}{\partial R_{11}} |R_{rs}| = \begin{vmatrix} R & 0 & \cdot & \cdot & 0 \\ 0 & R_3 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & R_n \end{vmatrix} \tag{57}$$

The quotient of (56) and (57) is simply R_1 and this quotient is invariant under the transformation (49). Thus if according to the last section several values of R_r are possible (e.g. $R_r R_\rho R_k$ in Fig. 16), the corresponding networks cannot be derivable from each other by a transformation like (49) although an infinite group can be derived from each one of them by such a transformation. This is true also when a usable value of $R(\Omega)$ occurs for $\Omega = 0$ or ∞. In such a case a network with a greater number of meshes results.

The question naturally presents itself whether all equivalent networks can be derived by the application of the transformation (49) to each of the networks obtained by using separate R_ρ.

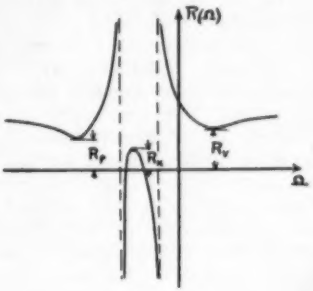


FIG. 16

This question must be answered in the negative, for examples can be constructed in which the quotient $\dfrac{|R_{rs}|}{\dfrac{\partial}{\partial R_{11}}|R_{rs}|}$ for a given network does not appear as a stationary value of the function $R(\omega^2) = \text{Re}\, Z(i\omega)$ derived from the network. It is thus evident that a gap still remains to be bridged in the theory of equivalent networks. The results of bridging this gap should have considerable mathematical as well as physical interest.

BIBLIOGRAPHY

1. R. Foster: "A Reactance Theorem," *Bell Syst. Techn. Journ.*, 1924, vol. 3, p. 259.
2. R. Foster: "Theorems on the Driving-point Impedance of Two-mesh Circuits," *Bell Syst. Techn. Journ.*, 1924, vol. 3, p. 651.
3. W. Cauer: "Die Verwirklichung von Wechselstromwiderständen vorgeschriebener Frequenzabhängigkeit," *Archiv für Elektrot.*, 1926–27, Bd. 17, p. 355.
4. T. C. Fry: "The Use of Continued Fractions in the Design of Electric Networks," *Bull. of Am. Math. Soc.*, 35, p. 463 (1929).
5. W. Cauer: "Ueber eine Klasse von Funktionen die die Stieltjesschen Kettenbrüche als Sonderfall enthält," Jahresbericht d. Deutschen Mathematiker Ver., 38, 1929, p. 63.

6. N. Howitt: "On the Invariant Impedance Function and its Associated Group of Networks," M. I. T. doctor's thesis, 1930.—"Group Theory and the Electric Circuit" Phys. Rev. Vol. 37 Ser. 2, Num. 12, 1931, p. 1583.
7. H. G. Baerwald: "Ein enfacher Beweis des Reaktanztheorems," *Elektrische Nachrichten-Technik*, 1930, Bd. 7, p. 331.
8. W. Cauer: "Ueber die Variablen eines Passiven Vierpols," *Sitzungsber. d. Preuss. Ak. d. Wissensch.*, 1927 (Dez.), p. 268.
"Vierpole," *Elektrische Nachrichten-Technik*, 1929, Bd. 6, p. 272.
9. W. Cauer: "Untersuchungen über ein Problem das drei positiv definite quadratische Formen mit Streckenkomplexe in Beziehung setzt," *Mathematische Annalen*, 1931 (to appear shortly).
10. Yuk-Wing Lee: "Synthesis of Electric Networks by Means of the Fourier Transforms of Laguerre Functions," M. I. T. doctor's thesis, June, 1930.
11. A. Hurwitz: "Ueber die Bedingungen unter welchen eine Gleichung nur Wurzeln mit negativen reëllen Theilen besitzt," *Math. Ann.*, 1895, Bd. 46, p. 273.
12. W. Cauer: "Ein Satz über zwei zusammenhängende Hurwitz'sche Polynome," *Sitzungsberichte d. Berliner Mathem. Gesellsch.*, 1928, XXVII. Jahrg., p. 25.
13. G. Pick: "Konforme Abbildung kreisförmiger Bereiche," *Math. Ann.*, 1915–16, Bd. 77, p. 1.
14. L. Bieberbach: *Lehrbuch der Funktionentheorie*, I. u. II. B. G. Teubner, Leipzig, 1927.
15. F. Klein: *Vorlesungen über Nicht-euklidische Geometrie.* (Die Grundlagen der Mathematischen Wissenschaften in Einzeldarstellungen, Band 26.) Julius Springer, Berlin, 1928.

(*Letter to the Editor by G. W. O. Howe*)

Dear Sir:

I have read with great interest the paper by Dr. H. B. Dwight on the CALCULATION OF RESISTANCES TO GROUND by the method which I introduced in 1914. I was always under the impression that the method gave very accurate results, but Dr. Dwight shows conclusively that in the case of a disc the error is as much as 8 per cent and there is no doubt that in the case of a square the error will be of the same order. It is important to know that the method can give errors of this magnitude in certain cases. To enable those who are interested in the subject of the calculation by this method either of capacity or of earth resistances to refer to the earlier papers on the subject, I append a list of the papers published between 1914 and 1920.

I am,

Yours faithfully,

G. W. O. HOWE [1]

1914. "The Capacity of Radio-Telegraph Antennae." *The Electrician*, vol. 73, pp. 829, 859, 906.

1915. "The Capacity of Aerials of the Umbrella Type." *The Electrician*, vol. 75, p. 870.

1915. "The Calculation of the Effective Resistance of Earth Plates." *The Electrician*, Vol. 76, p. 353.

1916. "The Calculation of the Capacity of Radio-Telegraph Antennae Including the Effects of Masts and Buildings." *The Electrician*, Vol. 77, pp. 761, 880.

1917. "The Capacity of an Inverted Cone and the Distribution of its Charge." *Journal of the Physical Society of London*, XXIX, p. 339.

1920. "Capacity of Rectangular Plates and a Formula for the Capacity of Aerials." *Radio Review*, Nov. 1920, p. 710.

[1] Dr. Howe is Professor of Electrical Engineering at the University of Glasgow. Ed. Note.

DIRAC EQUATIONS IN GENERAL RELATIVITY

1. Four Dimensional Theory

By J. A. Schouten[1]

Introduction. In his paper "Elektron und Gravitation I"[2] Weyl has shown that Dirac's wave theory and general relativity can be connected in a quite satisfactory way by introducing a covariant differentiation of spin-quantities, and that the electromagnetic potential-vector comes in nearly automatically if we consider ψ not as an ordinary vector but as a pseudo-vector only determined but for an arbitrary "Eichfaktor" of the form $e^{i\varphi}$. He restricts himself in this paper only to the so called "two component" theory but deals with the "four component" theory in the second edition of his book on Gruppentheorie und Quantenmechanik. In 1930[3] the author has shown that the cause of the appearance of Hermitian symmetry in this part of physics is to be found in the non-definiteness of the fundamental tensor in relativity, and that it is possible to establish the theory for Minkowskian space-time in a way that is invariant under Lorentz transformations as well as under linear transformations in spin-space. Naturally the matrices generally used in dealing with this subject don't occur in a theory built in this way but they can be introduced at every moment by specializing the coordinate system in spin-space. It seemed desirable to generalize this for general Riemannian space-time and to establish on this base a theory of linear displacement of spin-quantities connecting them

[1] Professor Schouten, of the Delft Technical University, gave a course of lectures at the Massachusetts Institute of Technology during the fall term of 1930.—Ed. note.

[2] Zeitschrift fur Physik 50 (1929) 330–352.

[3] Die Darstellung der Lorentzgruppe in der komplexen E_4 abgeleitet aus den Diracschen Zahlen, Proc. Amsterdam 38 (1930), 189–197.

PRINCETON LIBRARY

with general relativity. The lectures on the general theory of linear displacement and its application on spin-quantities which I had the honor to give from October until December 1930 at the Massachusetts Institute of Technology and from January 1931 until March 1931 at Princeton University, led to the composition of this paper.

After a first § concerning symbols and equations of Ricci-calculus used, in which only the part concerning generalized pseudo-quantities brings something entirely new, the second § deals with the theory of spin-quantities in Minkowskian space-time. This § is a revised and much improved new edition of the ideas laid down in the short note cited above.[3] The α^k appear as co-contravariant quantities in spin-space and because of this fact expressions as $\bar{\psi}\alpha^k\psi$ have no meaning at all. No fundamental tensor in spin-space is introduced (as most authors do in some way, implicitly or explicitly) but by means of the *always* existing bivectordensity ϵ it is possible to deduce from the α^k a system of covariant quantities ϵ^k such that $\bar{\psi}\epsilon^k\psi$ has a meaning.

New is the conclusion that the quantities ϵ^k are densities of weight -1 and the components ψ^e and ψ^C of the spin-vector ψ densities of weight $+\frac{1}{2}$. The third § deals with spin-quantities in Riemannian space-time and their covariant differentiation. The chief theorem here proved is that the covariant differentiation of contravariant spinvectors of weight $+\frac{1}{2}$ is uniquely determined by the covariant differentiation in space-time. This § contains a discussion on more general pseudo-spin-vectors and leads to the conclusion that by the influence of a gravitational field the components of the potential vector change from ordinary numbers into Dirac numbers. The fourth § deals with the Dirac equations. The simplest covariant equation of the second order that can be linearized is deduced. The most interesting result of this § is that the Riemannian character of space-time shows itself chiefly in diminishing mc^2 by $\frac{1}{8}K_0\hbar^2$. The last § contains the variational equations, obtained by starting with a scalar density that differs from that used by Weyl only by a term due to the

non-symmetrical character of the displacement in space time. In addition to the terms found by Weyl in the above cited second edition p. 194 we find a term due to the above mentioned non-symmetrical character.

In the second part of this paper will be shown how the mass-term can be brought in in a natural way by starting with a five-dimensional Riemannian manifold.

§ 1. Symbols and Equations of Ricci Calculus Used in this Paper

1. Systems of measuring vectors. We denote the general coordinates in Riemannian space-time by ξ^ν; $\alpha, \cdots, \omega = 1, 2, 3, 4$ (printed in italics), and the corresponding measuring vectors by $\overset{\nu}{e}_\lambda$, $\overset{\nu}{e}_\lambda$. This system we denote by (ν). At every point we introduce an arbitrary system of four mutually perpendicular contravariant unitvectors,[4] one time-like and three space-like, $\overset{\nu}{e}_i$; $h, i, j, k, l, m = 0, 1, 2, 3$ (printed vertically), and the reciprocal covariant vectors $\overset{k}{e}_\lambda$. This system is denoted by (k). The components of any geometrical object [5] with more than one component with respect to (ν) are denoted by a central letter and small *greek* indices, those with respect to (k) by the *same* central letter and small *latin* indices. So $v^\nu = (v^1, v^2, v^3, v^4)$ are the components of a vector with respect to (ν) and $v^k = (v^0, v^1, v^2, v^3)$ those with respect to (k). If the system (ν) or (k) is transformed by introducing new coordinates ξ^N; $A, \cdots, \Omega = \dot{1}, \dot{2}, \dot{3}, \dot{4}$ or new measuring vectors $\overset{\nu}{e}_I$; $H, I, J, K, L, M = \dot{0}, \dot{1}, \dot{2}, \dot{3}$, new alphabets are used and new series of signs for the fixed indices but the *central letter of a geometrical object is never changed.* A quantity has not only components belonging to (ν) or to (k) but also *intermediate* components with respect to (ν) for some

[4] A unitvector is a vector the square of whose length is $+1$ if it is space-like and -1 if it is time-like.

[5] A *geometrical object* is a system of components with a well defined mode of transformation, for example the components of v^ν of a vector or the parameters $\Gamma^\nu_{\lambda\mu}$ of a displacement; a *quantity* is a geometrical object whose components transform linearly, for example a vector or a density.

of the indices and to (k) for the other ones. The intermediate components A^{ν}_{i}, A^{k}_{λ} of the unit affinor connect the components of a vector or affinor (but not those of more general geometrical objects) with respect to (ν) and to (k), e.g.

$$v^{k} = A^{k}_{\nu} v^{\nu}; \qquad v^{\nu} = A^{\nu}_{k} v^{k}. \tag{1}$$

When dealing with quantities with only one component, such as densities, the components with respect to (ν) and (k) are distinguished where there is danger of ambiguity by putting an index in brackets over the central letter, i.e. $\overset{(\nu)}{\mathfrak{p}}$, $\overset{(k)}{\mathfrak{p}}$. In general it is not possible to find coordinates ξ^{k} such that the system (k) is a system of measuring vectors belonging to those coordinates. This is possible if and only if

$$\partial_{[\mu} \overset{k}{c}_{\lambda]} \triangleq 0, \quad \text{equivalent to} \quad \partial_{[\mu} A^{k}_{\lambda]} \triangleq 0.^{6} \tag{2}$$

We call this case the *holonomic* case, the other the *non-holonomic* one.[7] In the nonholonomic case, we write $(d\xi)^{k}$ instead of $d\xi^{k}$ for the components of the linear element.

In Euclidean space, the case of § 2, we take the system (k) parallel at all points. Then ξ^{k} are orthogonal Cartesian coordinates. The general curvilinear coordinates ξ^{ν} do not occur in that §.

2. Covariant differentiation in Riemannian space-time. The parameters of the general metric non-symmetrical displacement with respect to (ν) are denoted by $\Gamma^{\nu}_{\lambda\mu}$, those of the ordinary symmetrical displacement by $\overset{0}{\Gamma}{}^{\nu}_{\lambda\mu}$. $\overset{0}{\Gamma}{}^{\nu}_{\lambda\mu}$ is symmetrical in λ and μ. So $\overset{0}{\Gamma}{}^{\nu}_{\lambda\mu} = {}'\{^{\lambda\mu}_{\nu}\}$, the Christoffel-symbol, and the covariant differentiations are given by

[6] The sign $\triangleq$ is always used when an equation is not invariant under all transformations of coordinates or measuring vectors with respect to any index or indices. So in general an equation with $\triangleq$ holds only for the system or systems of reference with respect to which the equation is written down and does not stand any transformation at all. In special cases some restricted transformations are allowed: in the case under consideration for instance all transformations of (ν) to holonomic systems.

[7] J. A. Schouten, Über nicht-holonome Übertragungen in einer L_n, *Math. Zeitschr.* 30 (1929), 149–172, S 153. The notations in that paper are not yet consequent with respect to the use of the central letter.

$$(3)\qquad \left.\begin{aligned} a)\ \ \nabla_\mu v^\nu &= \partial_\mu v^\nu + \Gamma^\nu_{\lambda\mu} v^\lambda \\ b)\ \ \overset{0}{\nabla}_\mu v^\nu &= \partial_\mu v^\nu + \overset{0}{\Gamma}{}^\nu_{\lambda\mu} v^\lambda \end{aligned}\right\} \partial_\mu = \frac{\partial}{\partial \xi^\mu}.$$

The parameters with respect to (k) are written Γ^k_{ij} and $\overset{0}{\Gamma}{}^k_{ij}$ respectively and we have:

$$(4)\qquad \left.\begin{aligned} a)\ \ \nabla_i v^k &= \partial_i v^k + \Gamma^k_{ij} v^i \\ b)\ \ \overset{0}{\nabla}_i v^k &= \partial_i v^k + \overset{0}{\Gamma}{}^k_{ij} v^i \end{aligned}\right\} \partial_i = A^\mu_i \partial_\mu.$$

$\overset{0}{\Gamma}{}^k_{ij}$ is in general not symmetrical in i and j. $\overset{0}{\Gamma}{}^\nu_{\lambda\mu}$, $\Gamma^\nu_{\lambda\mu}$ and $\overset{0}{\Gamma}{}^k_{ij}$, Γ^k_{ij} are connected by the equations [8]

$$(5)\qquad \overset{0}{\Gamma}{}^k_{ij} = A^{\lambda\mu k}_{ij\nu}\, \overset{0}{\Gamma}{}^\nu_{\lambda\mu} + A^k_\lambda \partial_i A^\lambda_i = A^{\lambda\mu k}_{ij\nu}\, \overset{0}{\Gamma}{}^\nu_{\lambda\mu} - A^{\mu\lambda}_{ji}\, \partial_\mu A^k_\lambda.$$

$$(6)\qquad \Gamma^k_{ij} = A^{\lambda\mu k}_{ij\nu}\, \Gamma^\nu_{\lambda\mu} + A^k_\lambda \partial_i A^\lambda_i = A^{\lambda\mu k}_{ij\nu}\, \Gamma^\nu_{\lambda\mu} - A^{\mu\lambda}_{ji}\, \partial_\mu A^k_\lambda.$$

From (5) follows

$$(7)\qquad \overset{0}{\Gamma}{}^k_{ij} \overset{*}{=} - A^{\mu\lambda}_{ji}\, \partial_{[\mu} A^k_{\lambda]}.$$ [9]

Hence $\overset{0}{\Gamma}{}^k_{ij}$ is only symmetrical in i and j in the holonomic case. From (2) and (6) we deduce that the $\Gamma^\nu_{[\lambda\mu]}$ transform like the components of a quantity of the third degree for all transformations to holonomic systems. We denote this quantity by $S_{\lambda\mu}{}^{;\nu}$

$$(8)\qquad \Gamma^\nu_{[\lambda\mu]} \overset{*}{=} S_{\lambda\mu}{}^{;\nu}.$$

It is well known that

$$(9)\qquad \Gamma^\nu_{\lambda\mu} = \overset{0}{\Gamma}{}^\nu_{\lambda\mu} + S_{\lambda\mu}{}^{;\nu} - S^\nu{}_{.\lambda\mu} - S^\nu{}_{.\mu\lambda},$$

and from (5) and (6) follows that this equation holds for all transformations to holonomic or non-holonomic systems. It is remarkable that not only the alternating part of $\Gamma^\nu_{\lambda\mu}$ in $\lambda\mu$ but also the symmetric part depends on $S_{\lambda\mu}{}^{;\nu}$. From (6) it follows that

$$(10)\qquad \Gamma^k_{[ij]} = S_{ij}{}^{;k} - A^{\mu\lambda}_{ji}\, \partial_{[\mu} A^k_{\lambda]},$$

[8] Der Ricci-Kalkül, Springer 1923, S 73.

[9] Here only transformations of (ν) to *holonomic* systems are allowed since it is only for these systems and not for general systems of measuring vectors that $\overset{0}{\Gamma}{}^\nu_{[\lambda\mu]}$ is always zero.

and according to (7)

$$\Gamma^k_{(ij)} = \overset{0}{\Gamma}{}^k_{(ij)} + S_{ij}{}^k, \tag{11}$$

both equations holding for all transformations to holonomic or non-holonomic systems.

It is a well known fact that for the orthogonal systems (k)

$$\overset{0}{\Gamma}{}^{(k)}_j = 0; \qquad \overset{0}{\Gamma}{}^{ik}_j = g^{ia}\overset{0}{\Gamma}{}^k_{ij}, \tag{12}$$

$$\overset{0}{\Gamma}{}^i_{ij} = 0. \tag{13}$$

But these equations hold as well for the Γ^k_{ij} as follows from (9):

$$\Gamma^{(k)}_j = 0; \qquad \Gamma^{ik}_j = g^{ia}\Gamma^k_{ij}, \tag{14}$$

$$\Gamma^i_{ij} = 0, \tag{15}$$

the first one expressing that an orthogonal system of unit-vectors passes by pseudoparallel displacement into a system of the same kind, the second one that the displacement leaves the volume invariant.

3. Densities.[10] A density of weight $\mathfrak{k}$ is a quantity of only one component with the transformation

$$\overset{(N)}{\mathfrak{p}} = \Delta^{-\mathfrak{k}}\overset{(\nu)}{\mathfrak{p}}; \qquad \Delta = |A^N_\lambda|, \tag{16}$$

if the system (ν) is transformed into the holonomic or non-holonomic system (N), $(d\xi)^N = A^N_\lambda d\xi^\lambda$ being the components of the linear element with respect to this system. The covariant derivative of $\mathfrak{p}$ is

$$\nabla_\mu \overset{(\nu)}{\mathfrak{p}} = \partial_\mu \overset{(\nu)}{\mathfrak{p}} - \mathfrak{k}\overset{(\nu)}{\mathfrak{p}}\Gamma^\alpha_{\alpha\mu}d\xi^\mu, \tag{17}$$

hence for the system (k):

$$\nabla_i \overset{(k)}{\mathfrak{p}} = \partial_i \overset{(k)}{\mathfrak{p}}. \tag{18}$$

[10] A theory of densities and the more simple pseudoquantities is found in J. A. Schouten und V. Hlavaty, Zur theorie der allgemeinen linearen Übertragung, *Math. Zeitschr.* 30 (1929), 414–432.

In Riemannian geometry

$$\overset{(\nu)}{\mathfrak{g}} = |g_{\lambda\mu}| \tag{19}$$

is a density of weight + 2, its square root

$$\overset{(\nu)}{\mathfrak{e}} = \sqrt{\overset{(\nu)}{\mathfrak{g}}}, \tag{20}$$

a most frequently used density of weight + 1. The components of $\mathfrak{g}$ and $\mathfrak{e}$ with respect to (k) have the value 1, hence from (18) it follows that the covariant derivatives of $\mathfrak{g}$ and $\mathfrak{e}$ vanish. An affinordensity of weight $\mathfrak{k}$ is the product of a density of weight $\mathfrak{k}$ and an ordinary affinor. The covariant differential contains a term involving $\Gamma^{\alpha}_{\alpha\mu}$, e.g. for the vectordensity $\mathfrak{v}^{\nu}$ of weight $\mathfrak{k}$

$$\delta\mathfrak{v}^{\nu} = d\mathfrak{v}^{\nu} + \Gamma^{\nu}_{\lambda\mu}\mathfrak{v}^{\lambda}d\xi^{\mu} - \mathfrak{k}\mathfrak{v}^{\nu}\Gamma^{\alpha}_{\alpha\mu}d\xi^{\mu}. \tag{21}$$

If we write $\Lambda^{\nu}_{\lambda\mu}$ for the parameters of the covariant differentiation of a contravariant vectordensity of weight $\mathfrak{k}$ or a covariant one with weight $-\mathfrak{k}$, then

$$\Lambda^{\nu}_{\lambda\mu} = \Gamma^{\nu}_{\lambda\mu} - \mathfrak{k}\Gamma^{\alpha}_{\alpha\mu}. \tag{22}$$

Hence $\Lambda^{\alpha}_{\alpha\mu}$ is zero if and only if $\mathfrak{k} = 1/n$.

The equation

$$\epsilon^{1234} = 1 \tag{23}$$

defines a contravariant vectordensity of weight + 1, that has only components + 1, − 1 and 0 with respect to *all* systems of measuring vectors. Its ordinary differential and its covariant differential both vanish.

Another kind of density used hereafter is defined after Weyl by the equation

$$\overset{(N)}{\mathfrak{p}} = |\Delta|^{-\frac{1}{2}}\overset{(\nu)}{\mathfrak{p}}, \tag{16'}$$

instead of (16). The covariant derivative of $\mathfrak{p}$ is then

$$\nabla_{\mu}\overset{(\nu)}{\mathfrak{p}} = \partial_{\mu}\overset{(\nu)}{\mathfrak{p}} - \mathfrak{k}\overset{(\nu)}{\mathfrak{p}}\overset{r}{\Gamma}{}^{\alpha}_{\alpha\mu}d\xi^{\mu}, \tag{17'}$$

where $\overset{r}{\Gamma}{}^{\alpha}_{\alpha\mu}$ is the real part of $\Gamma^{\alpha}_{\alpha\mu}$. For $\mathfrak{k} = 1/n$ we get in this case

$$\Lambda^{\nu}_{\lambda\mu} = \Gamma^{\nu}_{\lambda\mu} - \frac{1}{n}\overset{r}{\Gamma}{}^{\alpha}_{\alpha\mu}. \tag{22'}$$

Hence in this case $\Lambda^{\alpha}_{\alpha\mu}$ is imaginary and equal to the imaginary part of $\Gamma^{\alpha}_{\alpha\mu}$ if and only if $\mathfrak{f} = 1/n$.

4. Pseudoquantities.[9] In many cases we have besides the coordinates ξ^{ν} another independent variable ξ^{0} with the transformation

$$'\xi^{0} = \tau(\xi^{\nu})\xi^{0}; \qquad \tau \neq 0. \tag{24}$$

A pseudoscalar of class $\mathfrak{f}$ is a quantity with only one component $\overset{(\nu)}{\mathfrak{p}}$ with the transformation

$$\overset{(\nu)}{\mathfrak{p}} = \overset{\mathfrak{f}}{\tau}\overset{(\nu)}{\mathfrak{p}}, \tag{25}$$

where $\overset{\mathfrak{f}}{\tau}$ denotes the $\mathfrak{f}$-th power of τ. The covariant derivative of $\mathfrak{p}$ has the form

$$\nabla_{\mu}\overset{(\nu)}{\mathfrak{p}} = \partial_{\mu}\overset{(\nu)}{\mathfrak{p}} - \mathfrak{f}\Gamma_{\mu}\overset{(\nu)}{\mathfrak{p}}, \tag{26}$$

where Γ_{μ} is a geometrical object with the transformation

$$\Gamma_{M} = A^{\mu}_{M}\Gamma_{\mu} + \partial_{M}\log\tau. \tag{27}$$

The Γ_{μ} must be given in order that the pseudoscalars have a covariant differential and can't in general be computed from the $\Gamma^{\nu}_{\lambda\mu}$, this being possible only if τ depends on the transformation of ξ^{ν} into ξ^{N}. A density is such a pseudoscalar, here $\tau = \Delta^{-1}$. A pseudoaffinor of class $\mathfrak{f}$ is the product of a pseudoscalar of class $\mathfrak{f}$ and an ordinary affinor. The covariant differential contains a term in Γ_{μ}, thus for the pseudovector of class $\mathfrak{f}$:

$$\delta\mathfrak{v}^{\nu} = d\mathfrak{v}^{\nu} + \Gamma^{\nu}_{\lambda\mu}\mathfrak{v}^{\lambda}d\xi^{\mu} - \mathfrak{f}\mathfrak{v}^{\nu}\Gamma_{\mu}d\xi^{\mu}. \tag{28}$$

Pseudoquantities of this kind occur in cases where there is any indetermination. For instance an affinor $h_{\lambda\mu}$ of rank 1 can be written in an infinite number of ways as a product of two vectors and can therefore be considered as the product $\mathfrak{v}_{\lambda}\mathfrak{w}_{\mu}$ of two vectors respectively of class $+1$ and -1. Every ordinary quantity of degree higher than one can be considered as a sum of products of two or more pseudovectors provided that the sum of all the class-numbers of these pseudovectors be zero.

The idea of pseudoquantities can be generalized by introducing instead of ξ^0 n independent variables η^ν subject to linear transformations of the form

$$\eta^N = \mathfrak{A}^N_\nu(\xi^\nu)\eta^\nu; \qquad |\mathfrak{A}^N_\nu| = \sigma \neq 0, \tag{29}$$

where the $\mathfrak{A}^N_\nu$ are (like τ) functions of the ξ^ν. More general contravariant pseudovectors of class $\mathfrak{k}$ can then be defined by the transformation formulae

$$\mathfrak{v}^N = {}^{\mathfrak{k}}\sigma^{-1}\,\mathfrak{A}^N_\nu\mathfrak{v}^\nu; \qquad \mathfrak{v}^\nu = {}^{-\mathfrak{k}+1}\sigma\,\mathfrak{A}^\nu_N\mathfrak{v}^N. \tag{30}$$

From these the definition-formulae of more general covariant pseudovectors and more general pseudoquantities of higher degree can be deduced in the ordinary way. $\mathfrak{A}^\nu_\lambda$ is the unity affinor considered as a more general pseudo-quantity of class $+1$ in ν and -1 in λ. By transvection of a more general contravariant pseudovector of class $\mathfrak{k}$ and a covariant one of class $\mathfrak{k}'$ we get an ordinary pseudoscalar of class $\mathfrak{k}+\mathfrak{k}'$. The covariant differentiation of a more general contravariant pseudovector of class 1 has the form

$$\nabla_\mu\mathfrak{v}^\nu = \partial_\mu\mathfrak{v}^\nu + \Pi^\nu_{\lambda\mu}\mathfrak{v}^\lambda, \tag{31}$$

where $\Pi^\nu_{\lambda\mu}$ is a geometrical object with the transformation

$$\Pi^N_{\Lambda M} = \mathfrak{A}^{N\lambda}_{\nu\Lambda}A^\mu_M\Pi^\nu_{\lambda\mu} - \mathfrak{A}^\alpha_\Lambda\partial_M\mathfrak{A}^N_\alpha. \tag{32}$$

In the general case the group of transformations $\mathfrak{A}^N_\nu$ is the linear homogeneous one but a special case arises if a subgroup is used. The group consisting of the identical transformation only leads to ordinary quantities:

$$\mathfrak{A}^N_\nu \overset{*}{=} A^N_\nu. \tag{33}$$

The group of magnifications

$$\mathfrak{A}^N_\nu \overset{*}{=} \tau A^N_\nu; \qquad \mathfrak{A}^\lambda_\Lambda \overset{*}{=} \tau^{-1}A^\lambda_\Lambda, \tag{34}$$

leads to ordinary pseudoquantities. If the orthogonal group with respect to $g_{\lambda\mu}$ is used:

$$\mathfrak{A}^{\lambda\alpha}_{\Lambda M}g_{\lambda\mu} \overset{*}{=} g_{\Lambda M}, \tag{35}$$

$g_{\lambda\mu}$ is invariant for all transformations of the group and $g_{\lambda\mu}$ can therefore be considered as a pseudoquantity of class -1 in both indices:

$$\mathfrak{g}_{\lambda\mu} \overset{*}{=} g_{\lambda\mu}. \tag{36}$$

In this case the parameters $\Pi^{\nu}_{\lambda\mu}$ of the connection with $\delta\mathfrak{g}_{\lambda\mu} = 0$ are subject to the conditions

$$\Pi^{\alpha}_{\lambda\mu}\mathfrak{g}_{\alpha\nu} + \Pi^{\alpha}_{\nu\mu}\mathfrak{g}_{\alpha\lambda} = \overset{0}{\Gamma}{}^{\alpha}_{\lambda\mu}g_{\alpha\nu} + \overset{0}{\Gamma}{}^{\alpha}_{\nu\mu}g_{\alpha\lambda}, \tag{37}$$

from which we get

$$g_{\alpha(\nu}(\Pi^{\alpha}_{\lambda)\mu} - \overset{0}{\Gamma}{}^{\alpha}_{\lambda)\mu}) \overset{*}{=} 0. \tag{38}$$

Now every expression $\zeta_{\nu\lambda\mu}$ alternating in $\lambda\nu$ can be written

$$\zeta_{\nu\lambda\mu} = \eta_{\lambda\mu\nu} - \eta_{\nu\mu\lambda} - \eta_{\nu\lambda\mu}; \qquad \eta_{\lambda\mu\nu} = \zeta_{\nu[\lambda\mu]}, \tag{39}$$

hence $\Pi^{\nu}_{\lambda\mu}$ can be written

$$\Pi^{\nu}_{\lambda\mu} = \overset{0}{\Gamma}{}^{\nu}_{\lambda\mu} + \Sigma_{\lambda\mu}{}^{\cdot\nu} - \Sigma^{\nu}{}_{\cdot\lambda\mu} - \Sigma^{\nu}{}_{\cdot\mu\lambda}, \tag{40}$$

where $\Sigma_{\lambda\mu}{}^{\cdot\nu}$ is a geometrical object alternating in $\lambda\mu$. Hence *the parameters of covariant differentiation are derived from those of ordinary quantities by changing* $S_{\lambda\mu}{}^{\nu}$ *into* $\Sigma_{\lambda\mu}{}^{\cdot\nu}$. From (32) and (40) follow the formulae of transformation of the $\Sigma_{\lambda\mu}{}^{\cdot\nu}$:

$$\Sigma_{\Lambda\mathrm{M}}{}^{\cdot\mathrm{N}} = \mathfrak{A}^{\mathrm{N}\lambda}_{\nu\Lambda}A^{\alpha}_{\mathrm{M}}\Sigma_{\lambda\alpha}{}^{\cdot\nu} + \mathfrak{A}^{\mathrm{N}}_{\alpha}\partial_{[\mathrm{M}}\mathfrak{A}^{\alpha}_{\Lambda]}, \tag{41}$$

from which is to be seen that $\Sigma_{\lambda\mu}{}^{\cdot\nu}$ is an ordinary quantity as far as transformations of the ξ^{ν} alone are concerned ($\mathfrak{A}^{\mathrm{N}}_{\nu} \overset{*}{=} A^{\mathrm{N}}_{\nu}$) but that it changes with transformations of the η^{ν}. *Hence working with more general pseudoquantities defined with respect to the orthogonal group comes to the same thing as working with a variable* $S_{\lambda\mu}{}^{\nu}$.

5. Hermitian quantities.[11] Associated with the coordinates ξ^{ν} in any n-dimensional manifold we consider their complex conjugates which we write $\bar{\xi}^{\mathrm{N}}$, $\mathrm{N} = \bar{1}, \cdots, \bar{n}$. Then we can define vectors, affinors and other geometrical objects starting with the $\bar{\xi}^{\mathrm{N}}$ just in the same way as ordinary objects are defined with

[11] A theory of Hermitian quantities is found in J. A. Schouten and D. v. Dantzig: Über unitäre Geometrie, *Math. Ann.* (1930), 319–346.

respect to the transformation of the ξ^r. We call them geometrical objects *of the second kind*. Obviously every geometrical object of the first kind is in one to one correspondence with the geometrical object of the second kind having the complex conjugate components, so these latter objects don't give anything new. But something new arises if we consider geometrical objects with some small indices referring to ξ^r and some big ones referring to $\bar{\xi}^N$. Such objects we call *Hermitian quantities*. To every Hermitian quantity, for instance to $P_{\lambda\dot{M}}^{\cdot\cdot N}$ belongs one and only one other Hermitian quantity $\bar{P}_{\Lambda M}^{\cdot\cdot r}$ with components that are complex conjugate to the components of the first one. This quantity we call the *conjugated one, it is given the same central letter with a bar*. A covariant Hermitian quantity $h_{\lambda M}$ is called *symmetrical* resp. *alternating* if the covariant equation

$$h_{\lambda M} = \bar{h}_{M\Lambda} \quad \text{or} \quad h_{\lambda M} = -\bar{h}_{M\Lambda} \tag{42}$$

respectively holds. If any variable vector x^r is given, every symmetrical Hermitian quantity is in one to one correspondence with an Hermitian form

$$h_{\lambda M}\, x^{\lambda} \bar{x}^{M}. \tag{43}$$

§ 2. Spin-Quantities belonging to Euclidean Space-Time

If we introduce the simplest form of the wave equation in a euclidean space-time with the orthogonal cartesian coordinates ξ^k, $k = 0, 1, 2, 3$:

$$g^{ij}\partial_i\partial_j\psi = \frac{m^2c^2}{\hbar^2}\psi;\ \partial_i = \frac{\partial}{\partial\xi^i};\ \xi^0 = ct,\ \xi^1 = x,\ \xi^2 = y,\ \xi^3 = z, \tag{44}$$

where

$$-g^{00} = g^{11} = g^{22} = g^{33} = +1;\ g^{ij} = 0,\ i \neq j;\ i, j = 0, 1, 2, 3, \tag{45}$$

and the hypercomplex numbers of Dirac α^k, $k = 0, 1, 2, 3, 4$, defined by the equations

$$\begin{aligned} &\alpha)\ \alpha^{(i}\alpha^{j)} = g^{ij};\quad i, j = 0, 1, 2, 3,\\ &\beta)\ \alpha^4 = \alpha^0\alpha^1\alpha^2\alpha^3,\\ &\gamma)\ \alpha^4\alpha^4 = -1,\\ &\delta)\ (\alpha^i\alpha^j)\alpha^k = \alpha^i(\alpha^j\alpha^k);\quad i, j, k = 0, 1, 2, 3, 4, \end{aligned} \tag{46}$$

then this equation can be written as

$$\left(\alpha^j \frac{h}{i}\partial_j + mc\alpha^4\right)^2 \psi = 0, \tag{47}$$

and we may consider the linear equation

$$\left(\alpha^j \frac{h}{i}\partial_j + mc\alpha^4\right)\psi = 0 \tag{48}$$

as the *linearization* of (44).

The 16 numbers $1; \alpha^k; \alpha^{ij} = -\alpha^{ji} = \alpha^i\alpha^j; i \neq j; \alpha^{ijk} = \alpha^i\alpha^j\alpha^k; i, j, k \neq$ (equivalent to $\alpha^{4i} = \alpha^4\alpha^i$); α^4; $i, j, k = 0, 1, 2, 3$, generate an associative hypercomplex number system, the *sedenion-system*, first discovered by Clifford. Every number of this system (that is a linear combination of the 16 numbers first introduced with arbitrary real or complex coefficients) is called a *sedenion*. Now the co-contravariant quantities of the second degree $P^C_{\cdot A}$, $A, B, C, \cdots = 1, \cdots, 4$ in an *auxiliary* E_4 also build an associative numbersystem with 16 linear independent numbers if we define addition and multiplication as follows:

$$\begin{aligned} &a)\ (P+Q)^C_{\cdot A} = P^C_{\cdot A} + Q^C_{\cdot A}, \\ &b)\ \ (PQ)^C_{\cdot A} = P^C_{\cdot B}Q^B_{\cdot A}, \end{aligned} \tag{49}$$

and it can easily be proved that there can be established (in an infinite number of ways) a one to one correspondence between these two systems. Hence we may represent the 16 numbers, 1, α^k, etc., by co-contravariant quantities of the second degree in the auxiliary E_4 and accordingly write them $\alpha^C_{\cdot A}$, $\alpha^{kC}_{\cdot\cdot A}$, $\alpha^{ijC}_{\cdot\cdot\cdot A}$, $\alpha^{ijkC}_{\cdot\cdot\cdot\cdot A}$, $\alpha^{4iC}_{\cdot\cdot\cdot A}$, $\alpha^{4C}_{\cdot\cdot A}$. It is remarkable that these quantities are not the most general ones, for it can be proved that the elementary divisors of $\alpha^{kC}_{\cdot\cdot A}$ are $\lambda - 1, \lambda - 1, \lambda + 1, \lambda + 1$ for $k = 1, 2, 3$ and $\lambda - i, \lambda - i, \lambda + i, \lambda + i$ for $k = 0, 4$. Sedenions of this kind we call *space-like* or *time-like unit* sedenions respectively. A sedenion that can be obtained from a unitsedenion by multiplication with an ordinary real or complex number has elementary divisors of the form $\lambda - a, \lambda - a, \lambda + a, \lambda + a$ and is called a *special sedenion*. α^{ij}, α^{ijk} and α^{4i}, $i, j, k = 0, 1, 2, 3$; $i, j, k \neq$, are unitsedenions as well.

Now we can consider the 16 expressions $\alpha^{C}_{\cdot A}$, $\alpha^{kC}_{\cdot\cdot A}$, $\alpha^{ijC}_{\cdot\cdot\cdot A}$, $\alpha^{ijkC}_{\cdot\cdot\cdot\cdot A}$, $\alpha^{ijklC}_{\cdot\cdot\cdot\cdot\cdot A}$; $i, j, k, l = 0, 1, 2, 3$; $i, j, k, l \neq$, as quantities of the second, third, fourth, fifth and sixth degree lying with 0, 1, 2, 3, 4 indices in space-time and with two indices in the auxiliary E_4. By means of them every scalar, vector, bivector, trivector or quadrivector in space-time is connected in a unique way with a sedenion. On the other hand these quantities determine a one to one correspondence between any sedenion and certain systems consisting each of a scalar, a vector, a bivector, a trivector and a quadrivector in space time.[13]

All quantities in the E_4 we call *spinors*, the vectors also *spinvectors* and the E_4 *spin-space*.

By the choice of another orthogonal cartesian coordinate system in space-time α^0, α^1, α^2, α^3 undergo a Lorentz transformation and α^4 is invariant if the transformation has determinant $+1$ and invariant but for a factor -1 if the determinant is -1. Now every special sedenion, having the elementary divisors $\lambda - a$, $\lambda - a$, $\lambda + a$, $\lambda + a$ can be written by a suitable choice of the coordinate system in the form

$$(50) \qquad -a\underset{1}{e}^{C}\underset{1}{e}_{A} - a\underset{2}{e}^{C}\underset{2}{e}_{A} + a\underset{3}{e}^{C}\underset{3}{e}_{A} + a\underset{4}{e}^{C}\underset{4}{e}_{A},$$

and hence determines a system of two planes in E_4, the plane of $\underset{1}{e}$ and $\underset{2}{e}$ and the plane of $\underset{3}{e}$ and $\underset{4}{e}$. α^4 is invariant and so the two planes E_2 and $\bar{E}_2$ belonging to α^4 *are invariant planes in* E_4. The unity affinors in these planes are obviously

$$(51) \qquad \underset{0}{\imath} = \tfrac{1}{2}(\alpha + i\alpha^4); \qquad \underset{0}{\bar{\imath}} = \tfrac{1}{2}(\alpha - i\alpha^4),$$

and we have the identities

$$(52) \qquad \begin{aligned} &a)\ \alpha^4 = -i\underset{0}{\imath} + i\underset{0}{\bar{\imath}}, \\ &b)\ \alpha = \underset{0}{\imath} + \underset{0}{\bar{\imath}}. \end{aligned}$$

[13] The geometric properties of sedenions will be dealt with fully in the second part of this paper. See also the author's papers Über die in der Wellengleichung verwendeten hypercomplexen Zahlen, Proc. Kon. Akad. v. Wet. Amsterdam 32 (1929), 105–108 and Die Darstellung der Lorentzgruppe in der komplexen E_2 abgeleitet aus den Diracschen Zahlen, Proc. Kon. Ak. v. Wet. Amsterdam 33 (1930), 190–197.

By means of the unity affinors $\underset{0}{\iota}$ and $\underset{0}{\bar{\iota}}$ every spinvector can be decomposed into two components lying respectively in E_2 and $\bar{E}_2$. E_2 and $\bar{E}_2$ being invariant we now choose the contravariant measuring vectors in $\bar{E}_4$ so that two of them lie in E_2 and two in $\bar{E}_2$ and call them respectively $\underset{\mathfrak{a}}{e}^{\epsilon}$; $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$, $\cdots = 1, 2$ and $\underset{\mathfrak{A}}{e}^{\mathfrak{E}}$; $\mathfrak{A}$, $\mathfrak{B}$, $\mathfrak{E}$, $\cdots = \bar{1}, \bar{2}$. The components of any contravariant spinvector ψ^C are then $\psi^{\epsilon} = (\psi^1, \psi^2)$: $\psi^{\mathfrak{E}} = (\psi^{\bar{1}}, \psi^{\bar{2}})$.

Every sedenion P can be decomposed into four parts:

$$P = \underset{0}{\iota} P \underset{0}{\iota} + \underset{0}{\iota} P \underset{0}{\bar{\iota}} + \underset{0}{\bar{\iota}} P \underset{0}{\iota} + \underset{0}{\bar{\iota}} P \underset{0}{\bar{\iota}}. \tag{53}$$

Applying this decomposition to α^k, $k = 0, 1, 2, 3$, by virtue of (46β) and (51) two of the parts vanish, the other ones being

$$\left.\begin{aligned} \beta^k &= \underset{0}{\iota}\alpha^k \underset{0}{\bar{\iota}} = \underset{0}{\iota}\alpha^k = \alpha^k \underset{0}{\bar{\iota}} \\ \bar{\beta}^k &= \underset{0}{\bar{\iota}}\alpha^k \underset{0}{\iota} = \underset{0}{\bar{\iota}}\alpha^k = \alpha^k \underset{0}{\iota} \end{aligned}\right\} \alpha^k = \beta^k + \bar{\beta}^k. \tag{54}$$

β^k lies with the upper index in E_2 and with the lower in $\bar{E}_2$ and $\bar{\beta}^k$ just inverse, so that we can write them in full $\beta^{k\epsilon}_{\cdot\cdot\mathfrak{A}}$ and $\bar{\beta}^{k\mathfrak{E}}_{\cdot\cdot\epsilon}$. According to (46$\alpha$) they are subject to the identities

$$\left\{\begin{aligned} \beta^{(i}\,\bar{\beta}^{j)} &= \underset{0}{\iota} g^{ij} \\ \bar{\beta}^{(i}\,\beta^{j)} &= \underset{0}{\bar{\iota}} g^{ij}. \end{aligned}\right. \tag{55}$$

For the products of two and three factors β we get from (46), (51) and (55)

$$\left.\begin{aligned} \underset{1}{\iota} &= \quad i\beta^{01} = -i\beta^{10} = -\beta^{22} = \beta^{22} \\ &\qquad\qquad = \underset{0}{\iota}\alpha^{01} = -\underset{0}{\iota}\alpha^{22} \\ \underset{1}{\bar{\iota}} &= -i\bar{\beta}^{01} = \quad i\bar{\beta}^{10} = -\bar{\beta}^{22} = \bar{\beta}^{22} \\ &\qquad\qquad = -\underset{0}{\bar{\iota}}\alpha^{01} = -\underset{0}{\bar{\iota}}\alpha^{22} \end{aligned}\right\} \text{ cycl. } 1, 2, 3, \tag{56}$$

$$\left.\begin{aligned} \beta^{023} &= -i\beta^1; \quad \beta^0\bar{\beta}^{01} = -\beta^1; \quad \beta^{122} = -i\beta^0; \\ &\qquad\qquad\qquad\qquad\quad \beta^1\beta^{01} = -\beta^0 \\ \bar{\beta}^{023} &= \quad i\bar{\beta}^1; \quad \bar{\beta}^0\beta^{01} = -\bar{\beta}^1; \quad \bar{\beta}^{122} = \quad i\bar{\beta}^0; \\ &\qquad\qquad\qquad\qquad\quad \bar{\beta}^1\beta^{01} = -\bar{\beta}^0 \end{aligned}\right\} \text{cycl. } 1, 2, 3. \tag{57}$$

Here β^{ij}, β^{ijk}, $\overline{\beta}^{ij}$, $\overline{\beta}^{ijk}$, stand for $\beta^i\overline{\beta}^j$, $\beta^i\overline{\beta}^j\beta^k$, $\overline{\beta}^i\beta^j$, $\overline{\beta}^i\beta^j\overline{\beta}^k$; $i, j, k \neq$; other products don't occur because every product of a β with a β or a $\overline{\beta}$ with a β is identically zero. The equation (56) is used to define the quantities $\underset{i}{\iota}$ and $\underset{i}{\bar{\iota}}$, $i = 1, 2, 3$. These quantities are sedenions lying totally in E_2 and $\overline{E}_2$ respectively, so they can be written in full $\underset{i}{\iota}{}^{\mathfrak{c}}_{\cdot\mathfrak{a}}$ and $\underset{i}{\bar{\iota}}{}^{\mathfrak{G}}_{\cdot\mathfrak{H}}$. With $\underset{0}{\iota}$ and $\underset{0}{\bar{\iota}}$ respectively they each build a quaternion system:

$$(58)\qquad \left.\begin{array}{l} \underset{1}{\iota}\,\underset{2}{\iota} = -\,\underset{2}{\iota}\,\underset{1}{\iota} = \underset{3}{\iota}; \qquad \underset{1}{\bar{\iota}}\,\underset{2}{\bar{\iota}} = -\,\underset{2}{\bar{\iota}}\,\underset{1}{\bar{\iota}} = \underset{3}{\bar{\iota}} \\ \underset{1}{\iota}\,\underset{1}{\iota} = -\,\underset{0}{\iota} \qquad ; \qquad \underset{1}{\bar{\iota}}\,\underset{1}{\bar{\iota}} = -\,\underset{0}{\bar{\iota}} \end{array}\right\} \text{cycl. } 1, 2, 3.$$

The quantity β^k establishes a one to one correspondence between vectors in time-space and sedenions lying with the upper index in E_2 and with the lower in $\overline{E}_2$:

$$(59)\qquad \begin{aligned} w^{\mathfrak{c}}_{\cdot\mathfrak{H}} &= w_i\overline{\beta}^{i\mathfrak{c}}_{\cdot\cdot\mathfrak{H}} \\ w_j &= \tfrac{1}{2}w^{\mathfrak{c}}_{\cdot\mathfrak{H}}\beta^{\mathfrak{G}\mathfrak{H}}_{\cdot\cdot\mathfrak{c}}g_{ij} \end{aligned}$$

and $\overline{\beta}^k$ does the same with E_2 and $\overline{E}_2$ interchanged.

The quantities β^{ij} and $\overline{\beta}^{ij}$, $i, j = 0, 1, 2, 3$ establish a one to one correspondence between bivectors in time-space and systems of a sedenion lying in E_2 and one lying in $\overline{E}_2$:

$$(60)\qquad \begin{aligned} v^{\mathfrak{c}}_{\cdot\mathfrak{a}} &= v_{ij}\beta^{ij\mathfrak{c}}_{\cdot\cdot\cdot\mathfrak{a}} = \underset{1}{\iota}{}^{\mathfrak{c}}_{\cdot\mathfrak{a}}(-iv_{01} - v_{23}) + \text{cycl. } 1, 2, 3, \\ v^{\mathfrak{G}}_{\cdot\mathfrak{H}} &= v_{ij}\overline{\beta}^{ij\mathfrak{G}}_{\cdot\cdot\cdot\mathfrak{H}} = \underset{1}{\bar{\iota}}{}^{\mathfrak{G}}_{\cdot\mathfrak{H}}(iv_{01} - v_{23}) + \text{cycl. } 1, 2, 3, \end{aligned}$$

whose scalars $v^{\mathfrak{a}}_{\cdot\mathfrak{a}}$ and $v^{\mathfrak{H}}_{\cdot\mathfrak{H}}$ both vanish because $\underset{i}{\iota}$ and $\underset{i}{\bar{\iota}}$, $i = 1, 2, 3$, have the elementary divisors $\lambda + i$, $\lambda - i$.

If

$$(61)\qquad \begin{aligned} v^{\mathfrak{c}}_{\cdot\mathfrak{a}} &= \overset{1}{v}\underset{1}{\iota}{}^{\mathfrak{c}}_{\cdot\mathfrak{a}}, \\ v^{\mathfrak{G}}_{\cdot\mathfrak{H}} &= \overset{1}{v'}\underset{1}{\bar{\iota}}{}^{\mathfrak{G}}_{\cdot\mathfrak{H}}, \end{aligned}$$

the corresponding bivector has the components

$$(62)\qquad \left.\begin{aligned} v_{01} &= i(\overset{1}{v} - \overset{1}{v'}) \\ v_{23} &= -\overset{1}{v} - \overset{1}{v'} \end{aligned}\right\} \text{cycl. } 1, 2, 3.$$

Hence there is a one to one correspondence between *real* bivectors in space-time and sedenions expressible in terms of $\underset{1}{\iota}$, $\underset{2}{\iota}$ and $\underset{3}{\iota}$ with complex coefficients at one hand and with sedenions expressible in terms of $\underset{1}{\bar{\iota}}$, $\underset{2}{\bar{\iota}}$ and $\underset{3}{\bar{\iota}}$ with the complex conjugate coefficients on the other hand. So there is a one to one correspondence too between complex linear transformations in E_2 and $\bar{E}_2$, the transformations being conjugate if they have complex conjugate coefficients in $\underset{0}{\iota}$, $\underset{1}{\iota}$, $\underset{2}{\iota}$, $\underset{3}{\iota}$ and $\underset{0}{\bar{\iota}}$, $\underset{1}{\bar{\iota}}$, $\underset{2}{\bar{\iota}}$, $\underset{3}{\bar{\iota}}$. But this correspondence is not a linear one: if for instance the transformations P and Q correspond with $\bar{P}$ and $\bar{Q}$, then $(a + bi)P + (c + di)Q$ corresponds with $(a - bi)\bar{P} + (c - di)\bar{Q}$.

The most general linear transformation of contravariant vectors in E_2 is obtained by transvecting the vector with a quaternion of the form

$$Q = \overset{0}{Q}\underset{0}{\iota} + \overset{1}{Q}\underset{1}{\iota} + \overset{2}{Q}\underset{2}{\iota} + \overset{3}{Q}\underset{3}{\iota}: \tag{63}$$

$$'v^c = Q^c_{.a}v^a. \tag{64}$$

The corresponding transformation of covariant vectors is

$$'w_a = w_c \overset{-1}{Q}{}^c_{.a}, \tag{65}$$

where

$$\overset{-1}{Q} = \frac{1}{\|Q\|}(\overset{0}{Q}\underset{0}{\iota} - \overset{1}{Q}\underset{1}{\iota} - \overset{2}{Q}\underset{2}{\iota} - \overset{3}{Q}\underset{3}{\iota}) \tag{66}$$

and $\|Q\|$ is the norm of Q, that is the sum of the squares of the coefficients. If $\|Q\| = 1$, the transformations are unimodular. Hence the most general linear homogeneous transformation that leaves E_2 and $\bar{E}_2$ invariant transforms α^k in the following way

$$'\alpha^k = Q\beta^k\overset{-1}{R} + R\bar{\beta}^k\overset{-1}{Q}, \tag{67}$$

Q and R being arbitrary quaternions in E_2 and $\bar{E}_2$ respectively whose norms $\|Q\|$ and $\|R\|$ are not necessarily unity. The following theorems are easily proved:

I. A transformation of the form (67) leaves $\alpha^i\alpha^j g_{ij}$ invariant if the norms of Q and R are equal.

II. A transformation of the form (67) is a *real* transformation if and only if $\overset{0}{Q}, \overset{1}{Q}, \overset{2}{Q}, \overset{3}{Q}$ are complex conjugated to $\overset{0}{R}, \overset{1}{R}, \overset{2}{R}, \overset{3}{R}$ and the norms of Q and R are equal and *real*.

III. The real transformations expressible in the form mentioned under II are only those Lorentz transformations with determinant $+1$ that can be created out of real rotations in space and "special" Lorentz transformations, that are transformations of the form

$$
\begin{aligned}
{}'x &= x \cos h\lambda t + t \sin h\lambda t, \\
{}'t &= x \sin h\lambda t + t \cos h\lambda t.
\end{aligned} \tag{68}
$$

We get the whole group of *all* real Lorentz formations by introducing a transformation $'t = -t$ and one of the form $'x = -x$ but not all transformations got in this way can be expressed in the form mentioned under II. In other words, the group of all real transformations leaving $\alpha^i\alpha^j g_{ij}$ invariant is not continuous but a so called "mixed" group, consisting of such transformations that can be reached in the group-manifold [13] in a continuous way starting from identity and others where this is not possible. Only the first kind of transformations, which obviously build a continuous group for themselves,[14] can be expressed in the form mentioned under II.

So far we have made no assumption at all concerning the coordinate systems in E_2 and $\overline{E}_2$. Now we choose them in such a way that the ι and $\bar{\iota}$ and in consequence also the β and $\overline{\beta}$ get complex conjugate components. To make this choice invariant, from now on only complex conjugate linear transformations in E_2 and $\overline{E}_2$ are allowed. The determinants of β and $\overline{\beta}$ are not

[13] The group-manifold is the manifold in which every "point" is a transformation of the group.

[14] This group is not equivalent to the group of all real Lorentz transformations with determinant $+1$, for example $'x = x$, $'y = y$, $'z = -z$, $'t = -t$ belongs to this latter group and cannot be reached continuously starting from identity.

invariant with these transformations, but from (56) it follows that we always have

$$(69) \qquad |\beta^0| = -|\beta^k|; \qquad |\bar{\beta}^0| = -|\beta^k|; \qquad k = 1, 2, 3.$$

So we can choose the coordinates in E_2 and $\bar{E}_2$ in such a way that

$$(70) \qquad |\beta^0| = -|\beta^k| = |\bar{\beta}^0| = -|\bar{\beta}^k| = +1; \qquad k = 1, 2, 3,$$

and in order to make this choice invariant we have to put another restriction on the transformations of coordinates in E_2 and $\bar{E}_2$: the determinant has to be *real*. So from now on *only those transformations of the coordinates in E_2 and $\bar{E}_2$ are allowed, that applied to β^k and $\bar{\beta}^k$ induce in space-time the continuous subgroup of the Lorentz group mentioned above.*

The transformations of coordinates in E_2 and $\bar{E}_2$ being now complex conjugate, we can identify the quantities in $\bar{E}_2$ with the quantities of the second kind in E_2 (cf. § 1). Then the $\bar{\imath}$ are the quantities of the second kind conjugate to the $\imath$ and $\beta^{\bar{k}c}_{\cdot\cdot\mathfrak{A}}$ and $\bar{\beta}^{\bar{k}\mathfrak{C}}_{\cdot\cdot a}$ are conjugate Hermitian quantities in the last two indices.

In E_2 no bivector is given, as the quantities introduced so far have no comitant [15] that is a bivector. But the bivector-densities $\mathfrak{E}^{ab}$, ϵ_{ab} of weight $+1$ and -1 respectively in E_2 defined by the equations

$$(71) \qquad \mathfrak{E}^{12} = +1; \qquad \epsilon_{12} = +1,$$

holding for *every* system of coordinates, and their conjugates $\bar{\mathfrak{E}}^{\mathfrak{AB}}$, $\bar{\epsilon}_{\mathfrak{AB}}$ in $\bar{E}_2$ are *given* quantities. We now prove the *Principal Theorem:*

The densities of weight -1

$$(72) \qquad \begin{aligned} \epsilon^{\bar{k}}_{\cdot b\mathfrak{A}} &= \beta^{\bar{k}c}_{\cdot\cdot\mathfrak{A}}\epsilon_{cb} \\ \bar{\epsilon}^{\bar{k}}_{\cdot\mathfrak{B}a} &= \beta^{\bar{k}\mathfrak{C}}_{\cdot\cdot a}\bar{\epsilon}_{\mathfrak{CB}} \end{aligned}$$

[15] The word comitant is used in the sense of Weitzenbock, a comitant of any system of quantities is a quantity whose components depend on the components of the quantities of the system in a way independent of the choice of the coordinates. So an invariant is a comitant that is a scalar.

are Hermitian symmetrical in the last two indices:

$$\epsilon^{k}_{\cdot\alpha\bar{\beta}} = \bar{\epsilon}^{k}_{\cdot\beta\bar{\alpha}}. \tag{73}$$

In fact, from (55) and (70) follows immediately that

$$\left.\begin{aligned} \beta^{\dot{s}1}_{\cdot\cdot\bar{1}} &= -\bar{\beta}^{\dot{s}\bar{2}}_{\cdot\cdot 2} \\ \beta^{\dot{s}2}_{\cdot\cdot\bar{2}} &= -\bar{\beta}^{\dot{s}\bar{1}}_{\cdot\cdot 1} \\ \beta^{\dot{s}1}_{\cdot\cdot\bar{2}} &= +\bar{\beta}^{\dot{s}\bar{1}}_{\cdot\cdot 2} \\ \beta^{\dot{s}2}_{\cdot\cdot\bar{1}} &= +\bar{\beta}^{\dot{s}\bar{2}}_{\cdot\cdot 1} \end{aligned}\right\} k = 0, 1, 2, 3, \tag{74}$$

hence, by transvection with $\epsilon_{\dot{a}\dot{b}}$

$$\left.\begin{aligned} \epsilon^{k}_{\cdot 2} &= \bar{\epsilon}^{k}_{\cdot\bar{1}2} \\ \epsilon^{k}_{\cdot 12} &= \bar{\epsilon}^{k}_{\cdot\bar{2}1} \\ \epsilon^{k}_{\cdot\bar{2}\bar{2}} &= \bar{\epsilon}^{k}_{\cdot\bar{1}\bar{2}} \\ \epsilon^{k}_{\cdot\bar{1}\bar{1}} &= \bar{\epsilon}^{k}_{\cdot\bar{2}\bar{1}} \end{aligned}\right\} k = 0, 1, 2, 3. \tag{75}$$

Here is the point *where Hermitian symmetry and* the so called *Hermitian forms* (transvection of a symmetrical Hermitian quantity of the second degree with a vector and its conjugate) *come into this part of physics.* From our deduction it follows at once that their appearance depends on the $-$ sign of g^{00} [16] and on the special values of the determinants of the β. If we start with a positive definite fundamental tensor, Hermitian symmetry does not come in automatically and can only be brought in in an artificial way.[17]

From (70) and (72) follows for the determinants

$$|\epsilon^0| = |\bar{\epsilon}^0| = +1; \qquad |\epsilon^k| = |\bar{\epsilon}^k| = -1; \qquad k = 1, 2, 3, \tag{76}$$

hence ϵ^0 and $\bar{\epsilon}^0$ are positive definite and ϵ^k and $\bar{\epsilon}^k$; $k = 1, 2, 3$

[16] The Hermitian matrices occurring in Norbert Wiener's generalized harmonic analysis, *Acta Mathematica* vol. 55 (1930), pp. 117 to 258, have the same source, they only come in because the equation of light waves is of the hyperbolical type.

[17] This may be a hint not to change i into it in order to make things "more symmetrical," in fact doing this disturbs symmetry instead of improving it. Compare Eddington, The interaction of electric charges, *Proc. Roy. Soc.* 126 (1930), 696–728. Where he takes the $+$ sign for g^{00} he has, in building conjugates, to discriminate between the cases where i appears explicitly and those where it appears implicitly, as in it, and has not to be reversed.

negative definite. ϵ and $\bar{\epsilon}$ *establish a one to one correspondence between Hermitian tensor-densities of weight* -1 *and vectors in space-time.* It is easily shown that time-like vectors correspond to positive definite Hermitian tensordensities and space-like to indefinite ones. A vector in the light cone corresponds to a Hermitian tensordensity with determinant zero.

So far no attempt is made to fix the coordinates in E_2 and $\bar{E}_2$ in order to get definite components for the quantities ι, $\bar{\iota}$, β, $\bar{\beta}$ and ϵ. Now for our further developments this is in no way necessary or useful and we will give here a possible choice of coordinates merely in order to enable the reader to compare the results with those of other authors who work with a fixed coordinate system in E_4. ϵ^0 being positive definite and ϵ^1, ϵ^2, ϵ^3 indefinite, the system of unit-vectors $\underset{i}{e}^r$ being chosen in any way, coordinates in E_2 and $\bar{E}_2$ can be transformed always by an *allowed* transformation in such a way that ϵ^0 gets the simple form

$$\epsilon^0_{.1\bar{1}} = \epsilon^0_{.2\bar{2}} = +1; \qquad \epsilon^0_{.1\bar{2}} = \epsilon^0_{.2\bar{1}} = 0, \tag{77}$$

and ϵ^1, ϵ^2, ϵ^3 the following simple forms

$$\begin{aligned}
\epsilon^1_{.1\bar{1}} &= \quad \epsilon^1_{.2\bar{2}} = \quad 0; \quad \epsilon^1_{.1\bar{2}} = \quad \epsilon^1_{.2\bar{1}} = -1,\\
\epsilon^2_{.1\bar{1}} &= \quad \epsilon^2_{.2\bar{2}} = \quad 0; \quad \epsilon^2_{.1\bar{2}} = -\epsilon^2_{.2\bar{1}} = +i,\\
\epsilon^3_{.1\bar{1}} &= -\epsilon^3_{.2\bar{2}} = -1; \quad \epsilon^3_{.1\bar{2}} = \quad \epsilon^3_{.2\bar{1}} = \quad 0.
\end{aligned} \tag{78}$$

Then the components of the β, that are not zero, are

$$\begin{aligned}
\beta^{01}_{..\bar{2}} &= -\beta^{02}_{..\bar{1}} = +1,\\
\beta^{11}_{..\bar{1}} &= -\beta^{12}_{..\bar{2}} = -1,\\
\beta^{21}_{..\bar{1}} &= \quad \beta^{22}_{..\bar{2}} = -i,\\
\beta^{31}_{..\bar{2}} &= \quad \beta^{32}_{..\bar{1}} = +1,
\end{aligned} \tag{79}$$

and the non-vanishing components of $\underset{0}{\iota}$, $\underset{1}{\iota}$, $\underset{2}{\iota}$ and $\underset{3}{\iota}$

$$\begin{aligned}
\underset{0}{\iota}{}^1_{.1} &= \quad \underset{0}{\iota}{}^2_{.2} = +1,\\
\underset{1}{\iota}{}^1_{.2} &= \quad \underset{1}{\iota}{}^2_{.1} = +i,\\
\underset{2}{\iota}{}^1_{.2} &= -\underset{2}{\iota}{}^2_{.1} = -1,\\
\underset{3}{\iota}{}^1_{.1} &= -\underset{2}{\iota}{}^2_{.2} = +i.
\end{aligned} \tag{80}$$

The matrices of the ι are but for a factor i identical with the matrices of Pauli.[18]

§ 3. Spin-Quantities belonging to Riemannian Space-Time

In Riemannian space-time all equations of § 2 are valid in every point, except the equations (44), (47) and (48) containing ∂_i. To every point of space-time belongs a local spin-space E_4. The components with small latin indices are now components with respect to the local system of orthogonal unit-vectors $\underset{0}{e}^r$, $\underset{1}{e}^r$, $\underset{2}{e}^r$, $\underset{3}{e}^r$ chosen at every point in an arbitrary way. If only one β is considered, the measuring vectors in the local spin-space may be chosen in such a way that the *ordinary* differentials of the components of β and $\overline{\beta}$ vanish. (If two different β are considered, for instance one for the electron and another for the proton, this is in general not possible for the components of both β.) Then they remain zero, 1° if we subject the coordinates in spin-space to general linear homogeneous transformations with real determinant in E_2 and the complex conjugate ones in $\overline{E}_2$ (the transformations allowed in § 2) and *at the same time* the unit-vectors to the *corresponding* real Lorentz transformation, 2° if we subject the unit-vectors or the coordinates in spin-space or both to "synchrone" Lorentz transformations respectively "synchrone" transformations of the kind allowed in § 2. Only in the

[18] Many authors speak of "matrices" and deal with matrices instead of precisely defined quantities. Now in geometry a matrix is only a scheme of components of a quantity with respect to any system of measuring vectors and one and the same matrix can belong to wholly different quantities. Hence such a scheme has no geometrical meaning at all if the transformation of the components is not given, in other words if the particular quantity to which it belongs is not given. For a quantity in geometry is given only by its components with respect to any system of reference *together with* their mode of transformation. The symmetry of a matrix for instance has only a geometrical sense if the matrix belongs to a covariant or a contravariant quantity and multiplication of matrices has no sense at all in the case of two covariant or two contravariant quantities. Using only matrices and neglecting the modes of transformation may be in some simple cases a suitable substitute for the theory of quantities that is behind it, but in every more complicated case it gives only a glimpse of it and spoils the "fine-structure" of the theory.

case 1° the components of β and $\bar{\beta}$ have definite values remaining constant during the investigation.

Covariant differentiation of spin-quantities can be chosen now in such a way that the *covariant* differentials of β and $\bar{\beta}$ vanish too. This can be done in various ways but we shall show that *with respect to contra and covariant spin-vector-densities of weight* $+ \frac{1}{2}$ *resp.* $- \frac{1}{2}$ (defined with the absolute value of the determinant of transformation) *one and only one covariant differentiation with this property is possible.* If we write for the parameters of the covariant differentiation of spin-vectors $\Lambda'^{\mathfrak{c}}_{\mathfrak{a}j}$ and $\Lambda'^{\mathfrak{G}}_{\mathfrak{A}j}$ then the equation

$$0 = \nabla_j \beta^{\mathfrak{kc}}_{\cdot\cdot\mathfrak{A}} = \partial_j \beta^{\mathfrak{kc}}_{\cdot\cdot\mathfrak{A}} + \Gamma^{\mathfrak{k}}_{\mathfrak{l}j} \beta^{\mathfrak{lc}}_{\cdot\cdot\mathfrak{A}} + \Lambda'^{\mathfrak{c}}_{\mathfrak{b}j} \beta^{\mathfrak{kb}}_{\cdot\cdot\mathfrak{A}} - \bar{\Lambda}'^{\mathfrak{B}}_{\mathfrak{A}j} \beta^{\mathfrak{kc}}_{\cdot\cdot\mathfrak{B}}, \tag{81}$$

leads to

$$\Gamma^{\mathfrak{k}}_{\mathfrak{l}j} \beta^{\mathfrak{lc}}_{\cdot\cdot\mathfrak{A}} = - \partial_j \beta^{\mathfrak{kc}}_{\cdot\cdot\mathfrak{A}} - \Lambda'^{\mathfrak{c}}_{\mathfrak{b}j} \beta^{\mathfrak{kb}}_{\cdot\cdot\mathfrak{A}} + \bar{\Lambda}'^{\mathfrak{B}}_{\mathfrak{A}j} \beta^{\mathfrak{kc}}_{\cdot\cdot\mathfrak{B}}. \tag{82}$$

Now we assume that the ordinary differentials of the components of $\beta^{\mathfrak{kc}}_{\cdot\cdot\mathfrak{A}}$ and $\bar{\beta}^{\mathfrak{bG}}_{\cdot\cdot\mathfrak{a}}$ vanish; then, using (55) we get

$$2\Gamma^{\mathfrak{a}}_{j} = - \Lambda'^{\mathfrak{c}}_{\mathfrak{b}j} \beta^{\mathfrak{kb}}_{\cdot\cdot\mathfrak{A}} \bar{\beta}^{\mathfrak{lA}}_{\cdot\cdot\mathfrak{c}} + \bar{\Lambda}'^{\mathfrak{B}}_{\mathfrak{A}j} \bar{\beta}^{\mathfrak{lA}}_{\cdot\cdot\mathfrak{c}} \beta^{\mathfrak{kc}}_{\cdot\cdot\mathfrak{B}}. \tag{83}$$

According to (15) $\Gamma^{\mathfrak{a}}_{\mathfrak{l}}$ vanishes identically, hence

$$0 = - \Lambda'^{\mathfrak{a}}_{\mathfrak{a}j} + \bar{\Lambda}'^{\mathfrak{A}}_{\mathfrak{A}j}. \tag{84}$$

Hence the imaginary part of $\Lambda'^{\mathfrak{a}}_{\mathfrak{a}j}$ vanishes. Obviously (84) is invariant under all transformations of spin-coordinates of the kind allowed in § 2. If we write $\Lambda^{\mathfrak{c}}_{\mathfrak{a}j}$ and $\bar{\Lambda}^{\mathfrak{G}}_{\mathfrak{A}j}$ for the parameters of the covariant differentiations of contra and covariant spin-vector-*densities* of weight $\frac{1}{2}$ and $- \frac{1}{2}$ respectively

$$\begin{aligned} \Lambda^{\mathfrak{c}}_{\mathfrak{a}j} &= \Lambda'^{\mathfrak{c}}_{\mathfrak{a}j} - \tfrac{1}{2} \iota^{\mathfrak{c}}_{\mathfrak{a}} \overset{\mathfrak{r}}{\Lambda}{}^{\mathfrak{b}}_{\mathfrak{b}j}, \\ \bar{\Lambda}^{\mathfrak{G}}_{\mathfrak{A}j} &= \Lambda'^{\mathfrak{G}}_{\mathfrak{A}j} - \tfrac{1}{2} \bar{\iota}^{\mathfrak{G}}_{\mathfrak{A}} \overset{\mathfrak{r}}{\Lambda}{}^{\mathfrak{B}}_{\mathfrak{B}j}, \end{aligned} \tag{85}$$

where $\overset{\mathfrak{r}}{\Lambda}{}^{\mathfrak{b}}_{\mathfrak{b}j}$ is the real part of $\Lambda^{\mathfrak{b}}_{\mathfrak{b}j}$, then by alternating the right side of (83) we get

$$2\Gamma^{\mathfrak{a}}_{j} = \beta^{\mathfrak{a}\mathfrak{a}}_{\cdots\mathfrak{c}} \Lambda^{\mathfrak{c}}_{\mathfrak{a}j} + \bar{\beta}^{\mathfrak{a}\mathfrak{A}}_{\cdots\mathfrak{G}} \bar{\Lambda}^{\mathfrak{G}}_{\mathfrak{A}j}, \tag{86}$$

or using (51):

$$
\begin{aligned}
2\Gamma^{01}_j &= -2\Gamma^{10}_j = \beta^{01\mathfrak{a}}_{\ldots\mathfrak{c}}\Lambda^{\mathfrak{c}}_{\mathfrak{a}j} + \bar{\beta}^{01\mathfrak{A}}_{\ldots\mathfrak{C}}\bar{\Lambda}^{\mathfrak{C}}_{\mathfrak{A}j},\\
2\Gamma^{23}_j &= -2\Gamma^{32}_j = -i\beta^{01\mathfrak{a}}_{\ldots\mathfrak{c}}\Lambda^{\mathfrak{c}}_{\mathfrak{a}j} + i\bar{\beta}^{01\mathfrak{A}}_{\ldots\mathfrak{C}}\bar{\Lambda}^{\mathfrak{C}}_{\mathfrak{A}j},
\end{aligned}
\tag{87}
$$

from which equation $\Lambda^{\mathfrak{c}}_{\mathfrak{a}j}$ and $\Lambda^{\mathfrak{C}}_{\mathfrak{A}j}$ can be computed. We get two different equivalent forms:

$$
\Lambda^{\mathfrak{c}}_{\mathfrak{a}j} = -\tfrac{1}{4}\Gamma_{\alpha\beta}\beta^{\alpha\beta\mathfrak{c}}_{\ldots\mathfrak{a}}; \qquad \bar{\Lambda}^{\mathfrak{C}}_{\mathfrak{A}j} = -\tfrac{1}{4}\Gamma_{\alpha\beta}\bar{\beta}^{\alpha\beta\mathfrak{C}}_{\ldots\mathfrak{A}},
\tag{88}
$$

where $\Gamma_{\alpha\beta} = g_{\beta\lambda}\Gamma^{\lambda}_{\alpha}$ and

$$
\Lambda^{\mathfrak{c}}_{\mathfrak{a}j} = 3i\Gamma^{[01}_{j}\beta^{23]\mathfrak{c}}_{\ldots\mathfrak{a}}; \qquad \bar{\Lambda}^{\mathfrak{C}}_{\mathfrak{A}j} = -3i\Gamma^{[01}_{j}\bar{\beta}^{23]\mathfrak{C}}_{\ldots\mathfrak{A}}.
\tag{89}
$$

It is easy to verify that the real parts of $\Lambda'^{\mathfrak{c}}_{\mathfrak{a}j}$ and $\Lambda'^{\mathfrak{A}}_{\mathfrak{A}j}$ are not determined by (81) because they have no influence on the values of the right side of (81). We could have derived the same result (88, 89) without mentioning $\Lambda'^{\mathfrak{c}}_{\mathfrak{a}j}$, by considering $\beta^{\ldots\mathfrak{c}}_{\ldots\mathfrak{a}}$ as a density of weight $+\frac{1}{2}$ in $\mathfrak{c}$ and $-\frac{1}{2}$ in $\mathfrak{a}$ which is obviously allowed.

The displacement of vectors in spin-space or of vector densities of another weight is *not* determined by the displacement in time-space. But as we have only to deal with densities of this special weight this is of no importance.[19]

A general vectordensity $\psi^{\mathfrak{C}}$ of weight $+\frac{1}{4}$ in spin-space is determined by its two components $\psi^{\mathfrak{c}}$ and $\psi^{\mathfrak{C}}$ of weight $+\frac{1}{2}$ in E_2 and $\bar{E}_2$. The conjugate vectordensity $\bar{\psi}^{\mathfrak{C}}$ of the second kind is in the same way connected with $\bar{\psi}^{\mathfrak{C}}$ and $\bar{\psi}^{\mathfrak{c}}$. ψ determines in time-space two real scalars, two real vectors and a real bivector which can be defined by

$$
\begin{aligned}
\bar{\psi}\eta\alpha\psi &= -(\bar{\psi}^{\mathfrak{a}}\epsilon_{\mathfrak{a}\mathfrak{b}}\psi^{\mathfrak{b}} + \bar{\psi}^{\mathfrak{A}}\bar{\epsilon}_{\mathfrak{A}\mathfrak{B}}\psi^{\mathfrak{B}}) = B + \bar{B},\\
\bar{\psi}\eta\alpha^k\psi &= \bar{\psi}^{\mathfrak{a}}\epsilon^{k}_{\ldots\mathfrak{a}\mathfrak{B}}\psi^{\mathfrak{B}} + \bar{\psi}^{\mathfrak{A}}\bar{\epsilon}^{k}_{\ldots\mathfrak{A}\mathfrak{b}}\psi^{\mathfrak{b}} = s^k,\\
\bar{\psi}\eta\alpha^{ij}\psi &= -\bar{\psi}^{\mathfrak{a}}\epsilon_{\mathfrak{a}\mathfrak{c}}\beta^{ij\mathfrak{c}}_{\ldots\mathfrak{b}}\psi^{\mathfrak{b}} - \bar{\psi}^{\mathfrak{A}}\bar{\epsilon}_{\mathfrak{A}\mathfrak{C}}\bar{\beta}^{ij\mathfrak{C}}_{\ldots\mathfrak{B}}\psi^{\mathfrak{B}} = B^{ij},\\
\bar{\psi}\eta\alpha^{k4}\psi &= i(\bar{\psi}^{\mathfrak{a}}\epsilon^{k}_{\ldots\mathfrak{a}\mathfrak{B}}\psi^{\mathfrak{B}} - \bar{\psi}^{\mathfrak{A}}\bar{\epsilon}^{k}_{\ldots\mathfrak{A}\mathfrak{b}}\psi^{\mathfrak{b}}) = is'^k,\\
\bar{\psi}\eta\alpha^{4}\psi &= i(\bar{\psi}^{\mathfrak{a}}\epsilon_{\mathfrak{a}\mathfrak{b}}\psi^{\mathfrak{b}} - \bar{\psi}^{\mathfrak{A}}\bar{\epsilon}_{\mathfrak{A}\mathfrak{B}}\psi^{\mathfrak{B}}) = i(B - \bar{B}),
\end{aligned}
\tag{90}
$$

[19] In his paper Geometrie der Diracschen Theorie des Elektrons, Zeitschrift für Physik 57 (1929), 261–277, V. Fock has tried to make use of the indetermination of the displacement of spin-vectors to introduce the electromagnetic potential vector. However the displacement of contravariant vector-densities of weight $+\frac{1}{2}$ being wholly determined and only these vector-densities playing a role, the idea of Weyl of replacing the potential vectors by pseudo-vectors of class $+1$ and -1 seems much better.

where η is the bivector in spin-space defined by

$$\eta = -\epsilon - \bar{\epsilon}. \tag{91}$$

Vice versa ψ is determined by s^k, s'^k and B but for transformations of the form

$$'\psi^\epsilon = e^{i\varphi}\psi^\epsilon; \qquad '\overline{\psi}^\epsilon = e^{-i\varphi}\overline{\psi}^\epsilon. \tag{92}$$

It is easily proved that the bivector B^{ij} is a comitant of s, s' and B and can be written

$$B^{ij} = \frac{B + \overline{B}}{2|B|^2} s^{[i}s'^{j]} + \frac{i(B - \overline{B})}{2|B|^2} s_k s'_l I^{klij}, \tag{93}$$

where I^{klij} is the quadrivector with the orthogonal components $I^{0123} = +1$ in time-space. s and s' are mutually perpendicular and the absolute values of their lengths are equal and B^{ij} is a bivector whose two blades can be chosen one in the plane of s and s' and one perpendicular to this plane.

If we look at ψ only as an artificial wave symbol and at s, s' and the scalar B as quantities with a physical signification then ψ^ϵ and $\psi^{\mathfrak{C}}$ have to be considered as ordinary pseudo-vector-densities of class $+1$ and weight $+\frac{1}{2}$ with $\tau = e^{i\varphi}$ and $\overline{\psi}^\epsilon$ and $\overline{\psi}^{\mathfrak{C}}$ as pseudo-vector-densities of class -1 and weight $+\frac{1}{2}$. But we may go one step further and consider ψ only determined by s^k and B. Then ψ^ϵ and $\overline{\psi}^\epsilon$ are determined but for linear transformations of the form

$$'\psi^\epsilon = \tau\pi^\epsilon_{.\mathfrak{a}}\psi^\epsilon; \qquad '\overline{\psi}^\epsilon = \sigma\overline{\pi}^\epsilon_{.\mathfrak{a}}\overline{\psi}^{\mathfrak{a}}, \tag{94}$$

where τ and σ are chosen in such a way that the determinants of π and $\overline{\pi}$ are $+1$ and τ, σ, π and $\overline{\pi}$ are only bound by the condition that s and B are invariant. This condition leads to the equations

$$h^{\mathfrak{a}\mathfrak{B}} = \psi^{\mathfrak{a}}\overline{\psi}^{\mathfrak{B}} + \overline{\psi}^{\mathfrak{a}}\psi^{\mathfrak{B}} = \tau\bar{\tau}\pi^{\mathfrak{a}}_{.\epsilon}\psi^\epsilon\overline{\pi}^{\mathfrak{B}}_{.\mathfrak{D}}\psi^{\mathfrak{D}} + \sigma\bar{\sigma}\pi^{\mathfrak{a}}_{.\epsilon}\overline{\psi}^\epsilon\overline{\pi}^{\mathfrak{B}}_{.\mathfrak{D}}\psi^{\mathfrak{D}}, \tag{95}$$

$$\psi^{[\mathfrak{a}}\overline{\psi}^{\mathfrak{b}]} = \tau\sigma\pi^{[\mathfrak{a}}_{.\epsilon}\psi^{|\epsilon|}\overline{\pi}^{\mathfrak{b}]}_{.\mathfrak{d}}\overline{\psi}^{\mathfrak{d}}. \tag{96}$$

Now ψ^ϵ and $\overline{\psi}^\epsilon$ having equal rights (95) and (96) must hold if π

and $\bar{\pi}$ change places. From this it follows at once that

$$\sigma = \frac{1}{\tau} = e^{i\varphi} = \bar{\tau}, \tag{97}$$

$$\pi^{\epsilon}_{.\alpha} = \bar{\pi}^{\epsilon}_{.\alpha}, \tag{98}$$

and that $\pi^{\epsilon}_{.\alpha}$ corresponds to the most general transformation leaving $h^{\alpha\bar{\mathfrak{B}}}$ invariant:

$$\pi^{\alpha}_{.\epsilon}\bar{\pi}^{\bar{\mathfrak{B}}}_{.\mathfrak{r}}h^{\epsilon\mathfrak{r}} = h^{\alpha\bar{\mathfrak{B}}}. \tag{99}$$

Hence, if we introduce along with the coordinates in E_2 the auxiliary variables η^{ϵ} with the transformation

$$\eta^{\epsilon'} = \tau\pi^{\epsilon'}_{.\alpha}\eta^{\alpha}, \tag{100}$$

ψ^{ϵ} and $\bar{\psi}^{\epsilon}$ can be considered as more general pseudo-vector-densities respectively of class $+1$ or -1 and weight $+\frac{1}{2}$. The tensor $h^{\alpha\bar{\mathfrak{B}}}$ is invariant under the transformation $\pi^{\epsilon'}_{.\alpha}$, hence there exists a Hermitian more general pseudo quantity $\mathfrak{h}^{\alpha\bar{\mathfrak{B}}}$ of weight $+1$ in both indices having the same components as $h^{\alpha\bar{\mathfrak{B}}}$ with respect to all systems of reference

$$\mathfrak{h}^{\alpha\bar{\mathfrak{B}}} \doteq h^{\alpha\bar{\mathfrak{B}}}. \tag{101}$$

From this follows for the parameters $\Pi^{\epsilon}_{\alpha\jmath}$ and $\bar{\Pi}^{\mathfrak{E}}_{\mathfrak{A}\jmath}$ of the displacement of pseudo-quantities:

$$\Pi^{\epsilon}_{\alpha\jmath}\mathfrak{h}_{\epsilon\bar{\mathfrak{A}}} + \bar{\Pi}^{\mathfrak{E}}_{\mathfrak{A}\jmath}\mathfrak{h}_{\alpha\bar{\mathfrak{E}}} \doteq \Lambda^{\epsilon}_{\alpha\jmath}h_{\epsilon\bar{\mathfrak{A}}} + \bar{\Lambda}^{\mathfrak{E}}_{\mathfrak{A}\jmath}h_{\alpha\bar{\mathfrak{E}}}. \tag{102}$$

Hence

$$\varphi_{\mathfrak{A}\alpha\jmath} = (\Pi^{\epsilon}_{\alpha\jmath} - \Lambda^{\epsilon}_{\alpha\jmath})h_{\epsilon\bar{\mathfrak{A}}} \tag{103}$$

is Hermitian alternating in $\mathfrak{a}$ and $\mathfrak{A}$. The most general parameters Π are therefore given by the formulae

$$\Pi^{\epsilon}_{\alpha\jmath} = \Lambda^{\epsilon}_{\alpha\jmath} + \frac{ie}{ch}\varphi^{\epsilon}_{\alpha\jmath}, \tag{104}$$

where the factor $\frac{e}{ch}$ is introduced only for convenience and $\varphi^{\epsilon}_{\alpha\jmath}$ is

a geometrical object only bound by the condition

$$\varphi^{\varepsilon}_{\alpha\beta} h_{\varepsilon\overline{\beta}} = \overline{\varphi^{\varepsilon}_{\beta\alpha}} h_{\varepsilon\overline{\alpha}}. \tag{105}$$

From this condition it follows immediately that $\varphi^{\varepsilon}_{\alpha\beta}$ is real. An easy calculation shows that the introduction of the remaining part of $\varphi^{\varepsilon}_{\alpha\beta}$:

$$\varphi'^{\varepsilon}_{\alpha\beta} = \varphi^{\varepsilon}_{\alpha\beta} - \iota^{\varepsilon}_{\alpha}\varphi_{\beta}; \qquad \varphi_{\beta} = \tfrac{1}{2}\varphi^{\varepsilon}_{\varepsilon\beta} \tag{106}$$

is equivalent to replacing $S_{ij}{}^{k}$ by

$$S_{ij}{}^{k} + \Sigma'_{ij}{}^{k} + \Sigma_{ij}s^{k}, \tag{107}$$

where $\Sigma'_{ij}{}^{k}$ and Σ_{ij} are alternating in i and j and lie totally in the R_3 perpendicular to s^k. We have here a special case of the more general pseudo-quantities introduced in § 2, they belong to the group of those Lorentz transformations that leave the direction of s^k invariant. The introduction of $\varphi'^{\varepsilon}_{\alpha\beta}$ being equivalent with the introduction of an $S_{ij}{}^{k}$ with variable components that are functions of the η^{ε} we proceed here only with the other part of $\varphi^{\varepsilon}_{\alpha\beta}$ that depends only on φ_β. This means that we return in this paper to the point of view of Weyl. φ_β is no vector, its components transform according to the equation

$$\varphi_{\mu} = A^{\iota}_{\mu}\varphi_{\iota} - \frac{h}{e}\partial_{\mu}\varphi, \tag{108}$$

from which it follows that its alternated ordinary differential quotient is a bivector

$$\partial_{[\omega}\varphi_{\mu]} = f_{\omega\mu}. \tag{109}$$

As the φ_λ undergo only real transformations we suppose φ_λ to be real, $f_{\mu\lambda}$ is then real too and φ_λ and $f_{\mu\lambda}$ can be interpreted following Weyl as potential vector and electromagnetic bivector. In the more general case $\varphi^{\varepsilon}_{\alpha\beta}$ can be regarded as a potential vector whose components are not ordinary numbers but Dirac numbers, that is quantities of the second degree in E_2. But it is remarkable that to obtain this result it is not necessary to introduce more general pseudo-vectors, the part of $\Lambda^{\varepsilon}_{\alpha\beta}$ dependent on $S_{ij}{}^{k}$:

$$-\tfrac{1}{4}\beta^{\iota\kappa\varepsilon}_{\ldots\alpha}(2S_{\iota\beta\kappa} - S_{\kappa\iota\beta}), \tag{110}$$

together with $\varphi_i, \iota_a^\varepsilon$ is a potential vector with Dirac numbers for components as well. The only difference between the more general and the restricted theory is that in the former these components are functions of the η^ε. In both cases the part of $\Lambda_{\alpha j}^{\varepsilon}$ dependent on the Riemannian character of time-space and independent of S_{ij}^{k} will have a similar effect so that we come to the conclusion that *by the influence of a gravitational field the components of the potential vector change from ordinary numbers into Dirac-numbers.*

§ 4. Dirac Equations in Riemannian Space-Time

In Riemannian space-time the non-invariant equation (44) cannot be the wave equation and the simplest possible generalization cannot be linearized, $\nabla_i \nabla_j \psi$ being not symmetrical in i and j. So we start from the other end with the simplest possible generalization of (48):

$$\left(\alpha^j \frac{h}{i} \nabla_j + \frac{mc}{i} \alpha^4 \right) \psi = 0, \tag{112}$$

considering ψ as an ordinary pseudo-vector-density of weight $+ \frac{1}{4}$ and class 1 in E_4. Then (112) becomes, written in full:

$$\beta_{\ldots\mathfrak{B}}^{\mathfrak{K}} \frac{h}{i} \nabla_i \psi^{\mathfrak{B}} - mc\psi^{\mathfrak{C}} = \beta_{\ldots\mathfrak{B}}^{\mathfrak{K}} \left\{ \frac{h}{i} (\partial_i \psi^{\mathfrak{B}} + \overline{\Lambda}_{\mathfrak{A}i}^{\mathfrak{B}} \psi^{\mathfrak{A}}) + \frac{e}{c} \varphi_i \psi^{\mathfrak{B}} \right\} - mc\psi^{\mathfrak{C}} = 0, \tag{113}$$

$$\overline{\beta}_{\ldots\mathfrak{b}}^{\mathfrak{A}} \frac{h}{i} \nabla_i \psi^{\mathfrak{b}} + mc\psi^{\mathfrak{A}} = \overline{\beta}_{\ldots\mathfrak{b}}^{\mathfrak{A}} \left\{ \frac{h}{i} (\partial_i \psi^{\mathfrak{b}} + \Lambda_{ai}^{\mathfrak{b}} \psi^{a}) + \frac{e}{c} \varphi_i \psi^{\mathfrak{b}} \right\} + mc\psi^{\mathfrak{A}} = 0. \tag{114}$$

Iterating the process (112) we get

$$- h^2 \beta_{\ldots\mathfrak{B}}^{\mathfrak{K}} \beta_{\ldots a}^{\mathfrak{B}} \nabla_i \nabla_i \psi^{a} + m^2 c^2 \psi^{\varepsilon} = 0, \tag{115}$$

$$- h^2 \overline{\beta}_{\ldots\mathfrak{b}}^{\mathfrak{A}} \beta_{\ldots\mathfrak{A}}^{\mathfrak{b}} \nabla_i \nabla_i \psi^{\mathfrak{A}} + m^2 c^2 \psi^{\mathfrak{C}} = 0, \tag{116}$$

and these equations *are therefore in Riemannian space-time the*

simplest possible generalizations of (44) admitting linearization. Writing (115) out we get

$$
\begin{aligned}
(117)\quad - h^2\beta^{\mathfrak{k}\mathfrak{k}}_{\ldots\mathfrak{B}}\overline{\beta}^{\mathfrak{l}\mathfrak{B}}_{\ldots\mathfrak{a}}\Big\{ & \partial_i\partial_i\psi^{\mathfrak{a}} + (\partial_i\Lambda^{\mathfrak{a}}_{\mathfrak{b}\mathfrak{l}})\psi^{\mathfrak{b}} + 2\Lambda^{\mathfrak{a}}_{\mathfrak{b}\mathfrak{l}}\partial_j\psi^{\mathfrak{b}} \\
& + \frac{ie}{ch}(\partial_i\varphi_i)\psi^{\mathfrak{a}} + 2\frac{ie}{ch}\varphi_{\mathfrak{l}}\partial_j\psi^{\mathfrak{a}} - \frac{e^2}{c^2h^2}\varphi_i\varphi_i\psi^{\mathfrak{a}} \\
& - \Gamma^{\mathfrak{k}}_{\mathfrak{l}\mathfrak{j}}\partial_i\psi^{\mathfrak{a}} - \Gamma^{\mathfrak{k}}_{\mathfrak{l}\mathfrak{j}}\Lambda^{\mathfrak{a}}_{\mathfrak{b}\mathfrak{k}}\psi^{\mathfrak{b}} - \frac{ie}{ch}\Gamma^{\mathfrak{k}}_{\mathfrak{l}\mathfrak{j}}\varphi_i\psi^{\mathfrak{a}} \\
& + \Lambda^{\mathfrak{a}}_{\mathfrak{b}\mathfrak{j}}\Lambda^{\mathfrak{b}}_{\mathfrak{l}\mathfrak{i}}\psi^{\mathfrak{b}} + 2\frac{ie}{ch}\Lambda^{\mathfrak{a}}_{\mathfrak{b}\,\mathfrak{l}}\varphi_{\mathfrak{j}}\psi^{\mathfrak{b}}\Big\} + m^2c^2\psi^{\mathfrak{c}} = 0.
\end{aligned}
$$

Now a great number of terms in this expression can be made zero at one point P of space-time (not at all points at the same time) by a suitable choice of the local systems (k). Indeed, if we start with the local system at P and construct the local system at every point Q of the neighborhood of P by moving the system at P along the geodesic PQ in a pseudoparallel way, than not only $\delta e^{\mathfrak{k}}_i$ vanishes in P for every direction of $d\xi^r$ but also $\delta\delta e^{\mathfrak{k}}_i$ and from

$$
(118)\qquad \delta e^{\mathfrak{k}}_i \doteq \Gamma^{\mathfrak{k}}_{\mathfrak{l}\mathfrak{j}}(d\xi)^j = 0,
$$

$$
\begin{aligned}
(119)\qquad \delta\delta e^{\mathfrak{k}}_i &= (d\xi)^l\nabla_l(d\xi)^j\nabla_j e^{\mathfrak{k}}_i \\
&\doteq (d\xi)^l(d\xi)^j\partial_l\Gamma^{\mathfrak{k}}_{\mathfrak{l}\mathfrak{j}},
\end{aligned}
$$

and (87) it follows that not only $\Gamma^{\mathfrak{k}}_{\mathfrak{l}\mathfrak{j}}$ and $\Lambda^{\mathfrak{c}}_{\mathfrak{a}\mathfrak{j}}$ vanish but also $\partial_{\mathfrak{l}}\Gamma^{\mathfrak{k}}_{|\mathfrak{l}\mathfrak{j}|}$ and $\partial_{\mathfrak{l}}\Lambda^{\mathfrak{c}}_{|\mathfrak{a}|\mathfrak{j}}$.

Hence (117) becomes at P:

$$
\begin{aligned}
(120)\quad - h^2\beta^{\mathfrak{k}\mathfrak{k}}_{\ldots\mathfrak{B}}\overline{\beta}^{\mathfrak{l}\mathfrak{B}}_{\ldots\mathfrak{a}}\Big\{ & \partial_i\partial_i\psi^{\mathfrak{a}} + (\partial_i\Lambda^{\mathfrak{a}}_{\mathfrak{b}\mathfrak{l}})\psi^{\mathfrak{b}} + \frac{ie}{ch}(\partial_i\varphi_i)\psi^{\mathfrak{a}} \\
& + \frac{2ie}{ch}\varphi_{\mathfrak{l}}\partial_j\psi^{\mathfrak{a}} - \frac{e^2}{c^2h^2}\varphi_i\varphi_i\psi^{\mathfrak{a}}\Big\} + m^2c^2\psi^{\mathfrak{c}} = 0,
\end{aligned}
$$

or

$$
\begin{aligned}
(121)\quad & h^2g^{ij}\partial_i\partial_i\psi^{\mathfrak{c}} - m^2c^2\psi^{\mathfrak{c}} \\
& = g^{ij}\left\{\frac{eh}{ci}\partial_{\mathfrak{l}}\varphi_{\mathfrak{j}} + 2\frac{eh}{ci}\varphi_{\mathfrak{l}}\partial_j + \frac{e^2}{c^2h^2}\varphi_i\varphi_i\right\}\psi^{\mathfrak{c}} \\
& \qquad + \beta^{\mathfrak{k}\mathfrak{c}}_{\ldots\mathfrak{a}}\left\{\frac{eh}{ci}\psi^{\mathfrak{a}}\partial_{\mathfrak{l}}\varphi_{\mathfrak{j}} - h^2\psi^{\mathfrak{b}}\partial_{\mathfrak{l}}\Lambda^{\mathfrak{a}}_{|\mathfrak{b}|\mathfrak{j}}\right\}.
\end{aligned}
$$

Now at P

$$-h^2\beta^{\mu\epsilon}_{\cdots a}\partial_{[\mu}\Lambda^{a}_{|\nu\rho]} = 3h^2(\underset{0}{K} + K')\iota^{\epsilon}_{a}, \tag{122}$$

where

$$K_0 = -\frac{1}{12}K = -\frac{1}{12}g^{ij}K_{ijl} = \frac{1}{6}g^{ij}\partial_{[l}\overset{0}{\Gamma}{}^{k}_{|ij|k} \tag{123}$$

is the curvature of time-space for the case $S_{\lambda\mu}{}^{\nu} = 0$ and K' depends only on $S_{\lambda\mu}{}^{\nu}$ and its first covariant derivative. So (121) reduces to

$$\begin{aligned} h^2g^{ij}\partial_i\partial_j\psi^{\epsilon} - (m^2c^2 - 3h^2(\underset{0}{K} + K'))\psi^{\epsilon} \\ = g^{ij}\left\{\frac{eh}{ci}\partial_{[i}\varphi_{0} + \frac{2eh}{ci}\varphi_{[i}\partial_{j]} + \frac{e^2}{c^2h^2}\varphi_i\varphi_j\right\}\psi^{\epsilon} \\ + \frac{eh}{c}\{(-f_{01} + if_{23})\iota^{\epsilon}_{1a} + \text{cycl. } 1, 2, 3\}. \end{aligned} \tag{124}$$

Both terms to the right are well known from the euclidean theory, the last one is the spin term. The only correction due to the Riemannian character of space time is the term containing $\underset{0}{K} + K'$. It is remarkable that the whole term containing m vanishes if $\underset{0}{K} + K' = \dfrac{3mc^2}{h^2}$. According to a theorem of Schur in the symmetrical case $\underset{0}{K}$ can only be a constant in a Riemannian manifold with constant curvature and such a manifold can be imbedded in an R_5, its radius of curvature being then equal to $\sqrt{3}$ times $\dfrac{h}{mc}$, the principal wave length of the Broglie wave. The equation of such a manifold would be

$$x^2 + y^2 + z^2 - c^2t^2 - u^2 = -\frac{3h^2}{m^2c^2}, \tag{125}$$

u being the fifth Cartesian coordinate in R_5.

§ 5. Variational Equations

By transvecting (113) with $\overline{\psi}^{b}\epsilon_{bc}$ we get

$$-\frac{h}{i}\epsilon^{l}{}_{\lambda\beta}\overline{\psi}^{b}\nabla_l\psi^{\beta} - mc\overline{\psi}^{b}\epsilon_{bc}\psi^{c} = 0, \tag{126}$$

and by transvecting (114) with $\overline{\psi}^{\mathfrak{B}}\tilde{\epsilon}_{\mathfrak{B}\mathfrak{C}}$

$$-\frac{h}{i}\epsilon^{j}_{\mathfrak{b}\mathfrak{B}}\overline{\psi}^{\mathfrak{B}}\nabla_j\psi^{\mathfrak{b}}+mc\overline{\psi}^{\mathfrak{B}}\tilde{\epsilon}_{\mathfrak{B}\mathfrak{C}}\psi^{\mathfrak{C}}=0. \tag{127}$$

The sum of the right hand sides of (126) and (127) and $-\frac{h}{i}s^jS_{ij}^{\ \ k}$ multiplied by $\overset{(\nu)}{\epsilon}$:

$$\begin{aligned}\overset{(\nu)}{\mathfrak{M}}{}^{i}_{k}=&-\frac{h}{i}\overset{(\nu)}{\epsilon}\epsilon^{j}_{\mathfrak{b}\mathfrak{B}}(\overline{\psi}^{\mathfrak{B}}\nabla_j\psi^{\mathfrak{b}}+\overline{\psi}^{\mathfrak{b}}\nabla_j\psi^{\mathfrak{B}})-mc\overset{(\nu)}{\epsilon}\overline{\psi}^{\mathfrak{b}}\epsilon_{\mathfrak{b}\mathfrak{c}}\psi^{\mathfrak{c}}\\&+mc\overset{(\nu)}{\epsilon}\overline{\psi}^{\mathfrak{B}}\tilde{\epsilon}_{\mathfrak{B}\mathfrak{C}}\psi^{\mathfrak{C}}-\frac{h}{i}\overset{(\nu)}{\epsilon}s^jS_{ij}^{\ \ k}\end{aligned} \tag{128}$$

is a density of weight + 1 whose imaginary part is, *independently of* (126) and (127), a divergence:

$$\begin{aligned}\overset{(\nu)}{\mathfrak{M}}{}^{i}_{k}-\overset{(\nu)}{\overline{\mathfrak{M}}}{}^{i}_{k}=&-\frac{h}{i}\overset{(\nu)}{\epsilon}\epsilon^{j}_{\mathfrak{b}\mathfrak{B}}(\psi^{\mathfrak{b}}\overline{\psi}^{\mathfrak{B}}+\psi^{\mathfrak{b}}\psi^{\mathfrak{B}})-2\frac{h}{i}\overset{(\nu)}{\epsilon}s^jS_{ij}^{\ \ k}\\=&-\frac{h}{i}\overset{(\nu)}{\epsilon}\nabla_j s_i-2\frac{h}{i}\overset{(\nu)}{\epsilon}s^jS_{ij}^{\ \ k}\overset{*}{=}-\frac{h}{i}\partial_\mu\mathfrak{G}^\mu.\end{aligned} \tag{129}$$

The last equation holds only for holonomic systems. Hence $\overset{(\nu)}{\mathfrak{M}}$ seems well fit to be used in the equation of variation

$$\delta\int\overset{(\nu)}{\mathfrak{M}}{}^{i}_{k}d\xi^1d\xi^2d\xi^3d\xi^4 \tag{130}$$

Weyl calls a quantity whose imaginary part is a divergence " praktisch reell."

The independent variation of $\psi^{\mathfrak{c}}$, $\overline{\psi}^{\mathfrak{c}}$, $\psi^{\mathfrak{C}}$ and $\overline{\psi}^{\mathfrak{C}}$ with the condition that s^k be invariant leads immediately to the equations of Dirac (113) and (114) and to their conjugates. These latter equations we get more easily starting with $\overset{(\nu)}{\overline{\mathfrak{M}}}$. The independent variation of φ_i (only occurring in $\nabla_j\psi^{\mathfrak{b}}$ and $\nabla_j\psi^{\mathfrak{B}}$) leads immediately to $\frac{e}{c}s^k$, the one term of the Maxwellian equation containing the current vector-density. The other part of the Maxwellian equation comes in after adding to $\overset{(\nu)}{\overline{\mathfrak{M}}}$ the density $\overset{(\nu)}{\epsilon}f_{\lambda\mu}f^{\lambda\mu}$ with a suitable constant numerical coefficient.

We will now look upon $\overset{(\nu)}{\mathfrak{M}}$ as a function of the measuring vectors e_i^ν and give them a variation δe_i^ν. The spin-vectors ψ^c and $\overline{\psi}^c$, the $\epsilon^k_{.\alpha\beta}$, $g_{\lambda\mu}$ and $\Gamma^\nu_{\lambda\mu}$ may be constant during the variation. But the Γ^k_{ij} and $\overset{(\nu)}{e}$ change. $\overset{(\nu)}{\mathfrak{M}}$ can be decomposed into five parts.

$$
\begin{aligned}
\overset{(\nu)}{\mathfrak{M}}' &= -\frac{h}{i}\overset{(\nu)}{e}\epsilon^j_{.\mathfrak{a}\mathfrak{b}}(\overline{\psi}^{\mathfrak{a}}A^\mu_j\partial_\mu\psi^{\mathfrak{b}} + \overline{\psi}^{\mathfrak{b}}\Lambda^\mu_j\partial_\mu\psi^{\mathfrak{a}}),\\
\overset{(\nu)}{\mathfrak{M}}'' &= -\frac{h}{i}\overset{(\nu)}{e}\epsilon^j_{.\mathfrak{a}\mathfrak{b}}(\overline{\psi}^{\mathfrak{a}}\Lambda^{\mathfrak{b}}_{\mathfrak{a}'}\psi^{\mathfrak{a}'} + \overline{\psi}^{\mathfrak{b}}\Lambda^{\mathfrak{a}}_{\mathfrak{a}'}\psi^{\mathfrak{a}'}),\\
\overset{(\nu)}{\mathfrak{M}}''' &= -\frac{e}{c}\overset{(\nu)}{e}s^i\varphi_\mu A^\mu_j,\\
\overset{(\nu)}{\mathfrak{M}}^{\mathrm{IV}} &= -\frac{h}{c}\overset{(\nu)}{e}s^iA^\mu_jS^{\dots\alpha}_{\alpha\mu},\\
\overset{(\nu)}{\mathfrak{M}}^{\mathrm{V}} &= -mc\overset{(\nu)}{e}(\overline{\psi}^{\mathfrak{b}}\epsilon_{\mathfrak{b}\mathfrak{c}}\psi^{\mathfrak{c}} - \overline{\psi}^{\mathfrak{a}}\epsilon_{\mathfrak{a}\mathfrak{b}}\psi^{\mathfrak{b}}).
\end{aligned}
\tag{131}
$$

The first part gives a variation

$$
\delta\overset{(\nu)}{\mathfrak{M}}' = -\frac{h}{i}\epsilon^j_{.\mathfrak{a}\mathfrak{b}}e^j_L(\overline{\psi}^{\mathfrak{a}}\partial_\mu\psi^{\mathfrak{b}} + \overline{\psi}^{\mathfrak{b}}\partial_\mu\psi^{\mathfrak{a}})\delta e^\mu_j,
\tag{132}
$$

where $e^\mu_j = \overset{(\nu)}{e}e^\mu_j$. The second part can be written

$$
\begin{aligned}
\overset{(\nu)}{\mathfrak{M}}'' &= -\frac{h}{i}\overset{\nu}{e}\epsilon^j_{.\mathfrak{a}\mathfrak{b}}\{-\tfrac{1}{4}\overline{\psi}^{\mathfrak{a}}\psi^{\mathfrak{a}}\Gamma_{\mathfrak{a}\mathfrak{a}'}\beta^{\mathfrak{a}\mathfrak{a}'\mathfrak{b}}_{\dots\mathfrak{a}} - \tfrac{1}{4}\overline{\psi}^{\mathfrak{b}}\psi^{\mathfrak{a}}\Gamma_{\mathfrak{a}\mathfrak{a}'}\beta^{\mathfrak{a}\mathfrak{a}'\mathfrak{a}}_{\dots\mathfrak{a}}\\
&= -\tfrac{1}{2}\frac{h}{i}\overset{(\nu)}{e}\Gamma^k_{ik}s^i + 6h\overset{(\nu)}{e}s^{[0}\Gamma^{123]}.
\end{aligned}
\tag{133}
$$

Now

$$
\Gamma^k_{ik} = A^\lambda_i\Gamma^\alpha_{\lambda\alpha} + \partial_\alpha A^\alpha_i,
\tag{134}
$$

hence the first part of $\delta\overset{(\nu)}{\mathfrak{M}}''$ is, omitting parts that vanish by partial integration:

$$
\delta\overset{(\nu)}{\mathfrak{M}}'' = -\tfrac{1}{2}\frac{h}{i}\overset{(\nu)}{e}s^i\Gamma^\alpha_{\mu\alpha}e^j_l\delta e^\mu_j + \tfrac{1}{2}\frac{h}{i}\overset{(\nu)}{e}(\partial_\mu s^i)e^j_l\delta e^\mu_j
\tag{135}
$$

$$+ \tfrac{1}{2}\frac{h}{i}(\partial_\mu \overset{(\nu)}{\dot{e}})s^l \overset{j}{e}_l \delta e^\mu_j - \tfrac{1}{2}\frac{h}{i}(\delta \overset{(\nu)}{\dot{e}})s^l \Gamma^\alpha_{\mu\alpha} \overset{j}{e}_l \delta e^\mu_j$$

$$= \tfrac{1}{2}\frac{h}{i}(\partial_\mu s^l)\overset{j}{e}_l \delta e^\mu_j + \frac{h}{i} s^l \overset{j}{e}_l S_{\cdot\alpha\mu}^{\cdot\cdot\alpha} \delta e^\mu_j.$$

The second term of this expression is annihilated by

$$\delta \overset{(\nu)}{\mathfrak{W}}{}^{\mathrm{IV}} = -\frac{h}{i} s^l \overset{j}{e}_l S_{\alpha\mu}^{\cdot\cdot\alpha} \delta e^\mu_j. \tag{136}$$

The first term gives added to $\delta \overset{(\nu)}{\mathfrak{W}}{}'$

$$\tfrac{1}{2}\frac{h}{i} \epsilon^{\cdot\cdot j}_{\cdot\mathfrak{b}\mathfrak{B}} \overset{j}{e}_l (\psi^{\mathfrak{b}} \partial_\mu \overline{\psi}^{\mathfrak{B}} - \overline{\psi}^{\mathfrak{B}} \partial_\mu \psi^{\mathfrak{b}} + \overline{\psi}^{\mathfrak{b}} \partial_\mu \psi^{\mathfrak{B}} - \psi^{\mathfrak{B}} \partial_\mu \overline{\psi}^{\mathfrak{b}}) \delta e^\mu_j. \tag{137}$$

We may expect that the other parts of $\overset{(\nu)}{\mathfrak{W}}$ will give just the terms necessary to replace in this expression all ∂_μ by ∇_μ. The missing terms containing φ_μ are given by $\delta \overset{(\nu)}{\mathfrak{W}}{}'''$:

$$\delta \overset{(\nu)}{\mathfrak{W}}{}''' = -\frac{e}{c} \epsilon^{\cdot\cdot j}_{\cdot\mathfrak{b}\mathfrak{B}} \overset{j}{e}_l \varphi_\mu (\psi^{\mathfrak{b}} \overline{\psi}^{\mathfrak{B}} + \overline{\psi}^{\mathfrak{b}} \psi^{\mathfrak{B}}) \delta e^\mu_j, \tag{138}$$

and the other missing terms can be written

$$-6h \overset{i}{e}_\mu s \Gamma_{\underset{[0}{12|l|}} \delta e^\mu_{\underset{3]}{}}. \tag{139}$$

In fact we find after a somewhat laborious calculation that the second part of $\delta \overset{(\nu)}{\mathfrak{W}}{}''$ is equal to

$$\delta \overset{(\nu)}{\mathfrak{W}}{}'' = -6h \overset{i}{e}_\mu s \Gamma_{\underset{[0}{12|l|}} \delta e^\mu_{\underset{3]}{}} - 6h \overset{(\nu)}{\dot{e}} (\nabla_\beta s_\alpha) A^{\alpha\mathfrak{B}\omega}_{[0123]} \overset{j}{e}_\omega \mathfrak{E}_{\lambda\mu} \delta e^\mu_j + 18h \overset{(\nu)}{\dot{e}} S_{\beta\alpha}^{\cdot\cdot[\gamma} s_\gamma A^{\lambda\omega]\alpha\beta}_{[0123]} \overset{j}{e}_\omega \mathfrak{E}_{\lambda\mu} \delta e^\mu_j. \tag{140}$$

The fifth part gives the variation

$$\delta \overset{(\nu)}{\mathfrak{W}}{}^{\mathrm{V}} = -mc(\overline{\psi}^{\mathfrak{b}} \epsilon_{\mathfrak{b}\mathfrak{c}} \psi^{\mathfrak{c}} - \psi^{\mathfrak{B}} \epsilon_{\mathfrak{B}\mathfrak{C}} \psi^{\mathfrak{C}}) \delta \overset{(\nu)}{\dot{e}}$$
$$= \tfrac{1}{2}mc(\overline{\psi}^{\mathfrak{b}} \epsilon_{\mathfrak{b}\mathfrak{c}} \psi^{\mathfrak{c}} - \overline{\psi}^{\mathfrak{B}} \epsilon_{\mathfrak{B}\mathfrak{C}} \psi^{\mathfrak{C}}) \overset{j}{e}_\mu \delta e^\mu_j. \tag{141}$$

So, if we write

$$\delta \int \overset{(v)}{\mathfrak{M}} d\xi^1 d\xi^2 d\xi^3 d\xi^4 = \int T_{ij} \overset{j}{e}{}^i A^i_\mu \delta e^\mu_j, \tag{142}$$

then we find

$$\begin{aligned} T_{ij} = {} & \tfrac{1}{2} \frac{h}{i} \epsilon_{\mathfrak{A}\mathfrak{B}} (\psi^{\mathfrak{b}} \nabla_i \overline{\psi}^{\mathfrak{B}} - \overline{\psi}^{\mathfrak{B}} \nabla_i \psi^{\mathfrak{b}} + \overline{\psi}^{\mathfrak{b}} \nabla_i \psi^{\mathfrak{B}} - \psi^{\mathfrak{B}} \nabla_i \overline{\psi}^{\mathfrak{b}}) \\ & - 6h \nabla_{[\beta} S_{\alpha]} A^{\alpha\beta\omega}_{[0123]} g_{\omega\beta} g_{\lambda t} + 18h S_{\beta\alpha}{}^{[\gamma} S_\gamma A^{\lambda\omega]\alpha\beta}_{[0123]} g_{\omega\beta} g_{\lambda t} \\ & + \tfrac{1}{3} mc (\overline{\psi}^{\mathfrak{b}} \epsilon_{\mathfrak{b}\mathfrak{c}} \psi^{\mathfrak{c}} - \overline{\psi}^{\mathfrak{B}} \epsilon_{\mathfrak{B}\mathfrak{C}} \psi^{\mathfrak{C}}) g_{ij}.^{20} \end{aligned} \tag{143}$$

The influence of $S_{\lambda\mu}{}^{\nu}$ is to be seen in the third term.

[20] Comp. Weyl. Gruppentheorie und Quantenmechanik. S. 194.

DIRAC EQUATIONS IN GENERAL RELATIVITY

2. Five Dimensional Theory

By J. A. Schouten [1]

Introduction

The chief object of this second part is to show how the mass-term in the Dirac equations comes in automatically if we start with a five-dimensional instead of with a four-dimensional Riemannian manifold. After a first § concerning conventions and notations we show in § 2 that local spin-space in a point of a V_5 is the same as in a point of a V_4, but that α^4 has no longer an exceptional position. The sedenions, mentioned in the first part of this paper, are here dealt with thoroughly. The two most important theorems of this section are that two arbitrary pentad-systems have always ∞^1 special sedenions in common and that to every pentad system belongs a bivectordensity of weight $+\frac{1}{2}$ in spin-space. From the first theorem results that two given pentads determine a direction in the local R_5 and two invariant planes in local spin-space and this leads automatically back to the four-dimensional theory. From the second theorem results the possibility to deduce from every co-contravariant quantity α^k a co-covariant one η^k, so that $\overline{\psi}\eta^k\psi$ has a meaning though $\overline{\psi}\alpha^k\psi$ has no meaning at all. The symmetry of the η^k is equivalent with the Hermitian symmetry of the ϵ^k and $\bar{\epsilon}^k$ in the four-dimensional theory, so Hermitian symmetry comes in only after the introduction of two different α's. In the third § the covariant differentiation of spin-quantities is deduced. The simplest wave equation of the first order, $\alpha^k\nabla_k\psi^C = 0$ differs slightly from the equation in the four-dimensional case and this difference gives

[1] Professor Schouten, of the Delft Technical University, gave a course of lectures at the Massachusetts Institute of Technology during the fall term of 1930.—Ed. note.

rise to a term that has just the properties of a mass-term if 1°. time-space is holonomic in V_5 and 2°. the covariant differential of the spin-vector ψ^C in a direction perpendicular to space-time vanishes. Realisation is possible in different ways, but the case where the curvature of space-time is constant and equal to $-\frac{mc^2}{4h^2}$ may have a special interest.

1. Conventions and Notations

We start in this paper with a Riemannian five-dimensional manifold V_5 with a fundamental tensor of signature $+1$. The general coordinates are denoted by ξ^ν, $\alpha, \cdots, \omega = 1, 2, 3, 4, 5$ (printed cursive) and the corresponding measuring vectors by $\underset{\lambda}{e^\nu}$, $\underset{\Lambda}{e^\nu}$. This system we denote by (ν). At every point we introduce in the local R_5 an arbitrary system of five mutually perpendicular real contravariant unit vectors, two time like (with square of length $= -1$) $\underset{0}{e^\nu}$, $\underset{4}{e^\nu}$ and three space-like (with square of length $= +1$) $\underset{1}{e^\nu}$, $\underset{2}{e^\nu}$, $\underset{3}{e^\nu}$. This system is denoted by (c). The indices $a, b, \cdots, g$ will take the values 0, 1, 2, 3, 4 (printed vertically), the indices $h, i, \cdots, m$ only the values 0, 1, 2, 3. The component of any geometrical object gets small greek indices as far as it is referred to (ν) and small latin indices as far as it is referred to (c). All conventions and notations as to the use of central letters, transformations of coordinates, quantities with one component, non-holomic linear elements and intermediate components are just the same as in the four-dimensional case and are dealt with in § 1 of the first part of this paper, referred to in this part as I. Everything said concerning covariant differentiation, densities and pseudoquantities in §§ 2, 3, 4 of I holds for the V_5 too.

2. Spin-Quantities Belonging to a Point of V_5

At every point of the V_5 we have

$$-g^{00} = g^{11} = g^{22} = g^{33} = -g^{44} = +1. \tag{1}$$

From this it follows that the equation

$$g^{ab}\partial_b\partial_a\psi = 0 \qquad (a, b = 0, 1, 2, 3, 4) \tag{2}$$

can be linearized:

$$\alpha^b\partial_b\psi = 0 \tag{3}$$

by introducing the hypercomplex number system defined by the equation

$$\alpha^{(a}\alpha^{b)} = g^{ab},$$
$$(\alpha^a\alpha^b)\alpha^c = \alpha^a(\alpha^b\alpha^c) \quad (a, b, c = 0, 1, 2, 3, 4). \tag{4}$$

These equations are equivalent to (46) of I, hence linearization is accomplished with the use of the same hypercomplex number system as is needed in the case of the V_4.[2] As we have seen in I this system leads to the introduction of an auxiliary E_4, the local *spin-space* at every point of the V_4, so in the V_5 the same thing has to be done and the only difference between local spin-space in V_4 and in V_5 is that in the latter case α^4 has no exceptional position and that for this reason there are no invariant planes in E_4.

As in I we call every sedenion with elementary divisors $\lambda - a$, $\lambda - a$, $\lambda + a$, $\lambda + a$ a *special sedenion* and if $a = 1$, or $a = i$ respectively, a space-like or a time-like *unit sedenion*. The five sedenions α^0, α^1, α^2, α^3, α^4 and every other system of 5 sedenions in this system having the same rules of multiplication we call a *pentad*,[3] all sedenions that can be expressed linearly with real or complex coefficients in the sedenions of a pentad are special sedenions and form the *pentad-system* belonging to this pentad. Obviously every pentad system contains ∞^{10} pentads and these pentads can be transformed into each other by orthogonal trans-

[2] This is a general rule, to V_{2p} and V_{2p+1} belongs the same system. The quaternion system belongs to V_3 and V_5.

[3] Eddington uses the word pentad only for the so called "four-point" sedenions viz. real sedenions that have the rules of multiplication $i\alpha^0$, α^1, α^2, α^3, $i\alpha^4$ and whose matrices are built up of 12 numbers 0, 2 numbers $+1$ and 2 numbers -1. We reject this restriction because it is not invariant under linear transformations in E_4 and any pentad being given, the measuring vectors in E_4 can always be chosen in such a way that $i\alpha^0$, α^1, α^2, α^3, $i\alpha^4$ have the "four-point" property.

formations with respect to g_{ab}. The pentad α^k can be considered as a quantity of the third degree $\alpha^{kC}_{\cdot\cdot A}$, $(A, B, C, \cdots = 1, 2, 3, 4)$ with one index in local five-dimensional space-time and two indices in local spin-space. This quantity establishes a one-to-one correspondence between vectors in space-time and special sedenions belonging to the pentad-system of α^k. Another pentad of the same system establishes another correspondence between these vectors and the same special sedenions, that can be got from the former one by an orthogonal transformation in local space-time. But a pentad of another pentad-system establishes a one-to-one correspondence between space-time vectors and a totally different system of special sedenions, the special sedenions belonging to this other pentad-system.

A unit sedenion is invariant under ∞^8 linear homogeneous transformations, as follows immediately from the form of the elementary divisors. Hence there are ∞^8 unit sedenions and ∞ special sedenions and in the E_{16} of all sedenions these special sedenions fill an X_9. Each pentad system fills an E_5 in this X_9, *hence two arbitrary pentad systems have* ∞^1 *special sedenions in common that differ from each other only by a real or complex numerical factor.*[4] Among these common special sedenions are 4 unit sedenions, 2 time-like and 2 space-like. From this we see that by two different arbitrary pentad systems (or pentads) in the local spin-space of any point of V_5 a direction in the local R_5 is uniquely determined. Two pentad fields determine a congruence in V_5.

If $\underset{1}{\alpha}{}^4 = \underset{2}{\alpha}{}^4$ is a common unit sedenion of the two pentads $\underset{1}{\alpha}$ and $\underset{2}{\alpha}$ and thus $\underset{4}{e}{}^r$ the unitvector in the direction determined in the local R_5, from the rules of multiplication of the α^c follows immediately that $\underset{2}{\alpha}$ can be written in the form

$$\underset{2}{\alpha}{}^k = \lambda\, \underset{1}{\alpha}{}^k + \mu\, \underset{1}{\alpha}{}^{k4}; \qquad \lambda^2 + \mu^2 = 1, \tag{5}$$

[4] Eddington has proved a special case of this theorem: From the 16 numbers $1, \alpha^k, \alpha^{ij}, \alpha^{ijk}, \alpha^4$ six pentads can be built, and any two of these pentads have a unit sedenion in common. "A Symmetrical Treatment of the Wave Equation," *Proc. Royal Soc.*, 121 (1928), p. 389.

hence:

Every Unit Sedenion Belongs to ∞^1 Pentad Systems.

The pentad systems having α^4 in common with the system α are connected in pairs, the one connected with α contains the pentad $\alpha'^k = \alpha^{k4}$, $\alpha'^4 = \alpha^4$. We call α' the pentad *coupled* to α (to avoid the expression conjugated) with regard to α^4. From (4) it follows that α is coupled to the pentad $-\alpha^{k4}$, α^4.

The E_5 of the pentad systems filling up an X_9 and two of them having always in common an E_1, *the total number of pentad systems is* ∞^5.

Be ϵ_{ABCD} the quadrivectordensity in E_4 of weight -1 defined by the invariant equations

$$\epsilon_{1234} = +1, \tag{6}$$

then for any pentad α

$$\alpha^{eE}_{\cdot\cdot A}\,\alpha^{eF}_{\cdot\cdot B}\,\alpha^{eG}_{\cdot\cdot C}\,\alpha^{eH}_{\cdot\cdot D}\,\epsilon_{EFGH} = \epsilon_{ABCD} \tag{7}$$

(not summed over e) and from this equation it follows that

$$\alpha^{eE}_{\cdot\cdot A}\,\alpha^{eF}_{\cdot\cdot B}\,\epsilon_{EFCD} = \epsilon_{ABGH}\,\alpha^{eG}_{\cdot\cdot C}\,\alpha^{eH}_{\cdot\cdot D} \tag{8}$$

(not summed over e). Hence the bivector-tensordensity of weight -1

$$\nu_{ABCD} = g_{ab}\,\alpha^{aE}_{\cdot\cdot A}\,\alpha^{bF}_{\cdot\cdot B}\,\epsilon_{EFCD} \tag{9}$$

is a comitant [5] of the pentad system of α. In the E_6 of all bivectors of E_4 this quantity is a covariant tensordensity of weight $-\frac{1}{3}$ and can be written $\nu_{\lambda\nu}$; $\lambda, \nu = 1, 2, 3, 4, 5, 6$. According to (9) its rank is 5. Hence if $E^{\nu_1\cdots\nu_6}$ is the 6-vector-density in E_6 of rank $+1$ defined by the covariant equation

$$E^{123456} = +1 \tag{10}$$

[5] A comitant (Komitante) of any given system of quantities is a quantity whose components can be calculated from the components of the given quantities only. Because there is no quadri-vector given in E_4 there does not exist any comitant bivector-tensor, only a bivector-tensordensity of weight -1.

then

$$\omega^{\lambda\rho} = \nu_{\lambda_1\rho_1\rho} \cdots \nu_{\lambda_\rho\rho_3} \mathrm{E}^{\lambda_1\cdots\lambda_6\lambda} \mathrm{E}^{\rho_1\cdots\rho_6\rho} \tag{11}$$

is a contravariant tensor density of weight $+ \frac{1}{3}$ and rank 1 in E_6. Such a density is always the product of two equal contravariant vectordensities of weight $+ \frac{1}{6}$ in E_6 and each of them is equivalent to a contravariant bivectordensity of weight $+ \frac{1}{2}$ in E_4. The bivectordensity that differs from the derived one only by a numerical factor and is normalized by means of the quadrivector density E^{ABCD}:

$$\mathrm{H}^{\eta[AB}\,\mathrm{H}^{\prime CD]} = \mathrm{E}^{ABCD}; \qquad \mathrm{E}^{1234} = +1, \tag{12}$$

we will denote by $\mathrm{H}^{\prime AB}$. $\mathrm{H}^{\prime AB}$ is a comitant of the pentad system of α but for its sign and the same holds for the covariant bivector density of weight $-\frac{1}{2}$ defined by the equation

$$\eta'_{CD} = \mathrm{H}^{\prime AB}\,\epsilon_{ABCD}. \tag{13}$$

Obviously η'_{AB} is normalized as well:

$$\eta'_{[AB}\,\eta'_{CD]} = \epsilon_{ABCD}. \tag{14}$$

Hence these quantities are but for a factor ± 1 invariant under all transformations that leave the pentad system invariant, that is under the group of all orthogonal transformations in R_6. It is inconvenient that η'_{AB} and $\mathrm{H}^{\prime AB}$ are only invariant but for a factor ± 1. But because they are invariant under identity, they are invariant under all transformations of the orthogonal group, that can be reached in the group manifold in a continuous way starting with identity. This part of the group is generated by the 10 infinitesimal transformations

$$\alpha^C_A + \tfrac{1}{2}\alpha^{abC}_{\ldots A}\,dt \tag{15}$$

leading to finite transformations of the form

$$\left.\begin{aligned} {}'\alpha^p &= \alpha^p \cos\lambda + \alpha^q \sin\lambda \\ {}'\alpha^q &= -\alpha^p \sin\lambda + \alpha^q \cos\lambda \end{aligned}\right\} p \neq q \tag{16}$$

for $p, q = 1, 2, 3$ or $p, q = 0, 4$ and

$$\left.\begin{aligned} {}'\alpha^p &= \alpha^p \cosh\lambda + \alpha^q \sinh\lambda \\ {}'\alpha^q &= \alpha^p \sinh\lambda + \alpha^q \cosh\lambda \end{aligned}\right\} p \neq q$$

for $p = 1, 2, 3; q = 0, 4$ or $p = 0, 4; q = 1, 2, 3$.

The condition of invariance of η'_{AB} under (15) is

$$\alpha^{abC}_{\ldots A}\,\eta'_{CB} + \alpha^{abD}_{\ldots B}\,\eta'_{AD} = 0, \tag{18}$$

hence the quantity

$$\eta'_{BC}\,\alpha^{abC}_{\ldots A} = \eta'_{AC}\,\alpha^{abC}_{\ldots B} \tag{19}$$

is *symmetrical in A and B*. From (18) it follows that

$$\alpha^{pqC}_{\ldots A}\,\alpha^{pqD}_{\ldots B}\,\eta'_{CD} = \left\{\begin{array}{ll} -\eta'_{AB} & \left\{\begin{array}{l} p = 0, 4; \quad q = 1, 2, 3 \\ p = 1, 2, 3; q = 0, 4 \end{array}\right. \\ +\eta'_{AB} & \left\{\begin{array}{l} p, q = 0, 4 \\ p, q = 1, 2, 3 \end{array}\right. \end{array}\right\} \text{not summed over } p, q. \tag{20}$$

Transvection of (20) with $\alpha^{rsE}_{\ldots C}\,\alpha^{rsF}_{\ldots D}$ and alternation over $pqrs$ gives

$$\alpha^{rC}_{\ldots A}\,\alpha^{rD}_{\ldots B}\,\eta'_{CD} = \left\{\begin{array}{ll} -\eta'_{AB}, & r = 0, 4 \\ +\eta'_{AB}, & r = 1, 2, 3 \end{array}\right\} \text{not summed over } r \tag{21}$$

from which equation it follows that the quantity

$$\eta'_{BC}\,\alpha^{cC}_{\ldots A} = -\eta'_{AC}\,\alpha^{cC}_{\ldots B} \tag{22}$$

is *alternating in A and B*.

Every bivector in E_4 can be decomposed in ∞^4 ways into two leaves. If we choose in E_4 a fixed E_3 not passing through the origin, the bivector is equivalent to a null system in this E_3 and the decomposition of the bivector is equivalent to the decomposition of the system of sliding vectors belonging to this null system into two sliding vectors. Every special sedenion determines two planes in E_4 and the special sedenions belonging to the pentad system α determine therefore ∞^4 systems of two planes. It is easily shown that these systems coincide with the ∞^4

systems of two blades in which the bivector η'_{AB} can be decomposed. From this we see that not only η'_{AB} is determined by α but that vice versa to every non degenerate bivector in E_4(or null system in the E_3) one and only one pentad system belongs. The theorem that two pentad systems have always ∞^1 special sedenions in common is equivalent to the well known theorem that two systems of sliding vectors can be decomposed simultaneously in one and only one way into two sliding vectors so that the lines of these vectors coincide. The planes through the origin and their lines are the planes of common special sedenions.

We now take two pentad systems α and α' and suppose that $\alpha^4 = \alpha'^4$ is the common unit sedenion and that α' is coupled to α

$$\begin{gathered}\alpha'^0 = \alpha^{04};\ \alpha'^1 = \alpha^{14};\ \alpha'^{01} = \alpha^{01};\ \alpha'^{23} = \alpha^{23} \\ \alpha^0 = -\alpha'^{04};\ \alpha^1 = -\alpha'^{14}.\end{gathered} \tag{23}$$

Let η_{AB} be the normalized bivectordensity belonging to the pentad system of α', then from the equations corresponding to (20) for α' it follows that

$$\alpha^{rC}_{\cdot\cdot A}\,\alpha^{rD}_{\cdot\cdot B}\,\eta_{CD} = \left\{\begin{array}{ll} \eta_{AB}; & r = 0, 4 \\ -\eta_{AB}; & r = 1, 2, 3, \end{array}\right\} \text{ not summed over } r \tag{24}$$

hence

$$\alpha^{cE}_{\cdot\cdot A}\,\alpha^{cF}_{\cdot\cdot B}\,\alpha^{cG}_{\cdot\cdot C}\,\alpha^{cH}_{\cdot\cdot D}\,\eta_{[EF}\,\eta'_{GH]} = -\eta_{[EF}\,\eta'_{GH]}, \tag{25}$$

(not summed over c) which is only possible if $\eta_{[AB}\,\eta'_{CD]}$ vanishes. Hence the two null systems belonging to η_{AB} and η'_{AB} are in involution, that is there exists a coordinate system in E_4 such that the non-vanishing components of η and η' are $\eta'_{12} = -1$, $\eta'_{34} = +1$, $\eta_{12} = -1$, $\eta_{34} = -1$. From (19) and (24) we deduce that

$$\eta^{pq}_{\cdot\cdot BA} = \eta_{BC}\,\alpha^{pqC}_{\cdot\cdot\cdot A} = \left\{\begin{array}{ll} \eta^{pq}_{\cdot\cdot AB}; & p, q = 0, 1, 2, 3, \\ -\eta^{pq}_{\cdot\cdot AB} & \left\{\begin{array}{l} p = 0, 1, 2, 3; q = 4, \\ p = 4; q = 0, 1, 2, 3, \end{array}\right. \end{array}\right. \tag{26a}$$

$$\eta^{r}_{\cdot BA} = \eta_{BC}\,\alpha^{rC}_{\cdot\cdot A} = \left\{\begin{array}{ll} +\eta^{r}_{\cdot AB}; & r = 0, 1, 2, 3, \\ -\eta^{r}_{\cdot BA}; & r = 4. \end{array}\right. \tag{26b}$$

If we now decompose all quantities with respect to the two planes of α^4 and use the coordinate systems (c) and ($\mathfrak{C}$) in these planes introduced in I, then $-\eta$ splits up into ϵ and $\bar{\epsilon}$, α^k into β^k and $\bar{\beta}^k$, η^k into ϵ^k and $\bar{\epsilon}^k$, $k = 0, 1, 2, 3$, and α^4 into $-i\,\underset{0}{\iota}$ and $+i\,\underset{0}{\bar{\iota}}$. The symmetry of $\eta^k_{.AB}$, $k = 0, 1, 2, 3$ is then equivalent with the Hermitian symmetry of ϵ:

$$\epsilon^k_{.\mathfrak{a}\mathfrak{B}} = \bar{\epsilon}^k_{.\mathfrak{B}\mathfrak{a}}; \qquad k = 0, 1, 2, 3. \tag{27}$$

Covariant Differentiation of Spin Quantities

Any pseudo-vectordensity ψ^C of weight $+\frac{1}{4}$ and class $+1$ with $\tau = e^{i\varphi}$ has a covariant derivative of the form

$$\nabla_b \psi^C = \partial_b \psi^C + \Lambda^C_{Ab}\psi^A + \frac{ei}{ch}\varphi_b \psi^C. \tag{28}$$

The Λ^C_{Aj} can be chosen in one and only one way so that the covariant derivative of $\alpha^{cC}_{\cdot\cdot A}$ vanishes. Obviously $\alpha^{cC}_{\cdot\cdot A}$ can be considered as a quantity of weight $+\frac{1}{4}$ and class $+1$ in C and weight $-\frac{1}{4}$ and class -1 in A. Then from the equation

$$0 = \partial_b \alpha^{cC}_{\cdot\cdot A} + \Gamma^c_{ab}\alpha^{aC}_{\cdot\cdot A} + \Lambda^C_{Bb}\alpha^{cB}_{\cdot\cdot A} - \Lambda^B_{Ab}\alpha^{cC}_{\cdot\cdot B} \tag{29}$$

and the identity

$$\Lambda^A_{Aj} = 0, \tag{30}$$

depending on the fact that ψ^C is of weight $+\frac{1}{4}$, we get

$$4\Gamma^{ac}_b + \alpha^{[a|A|}_{\cdot\cdot C}\,\partial_b\,\alpha^{c]C}_{\cdot\cdot A} = 2\Lambda^C_{Bb}\,\alpha^{acB}_{\cdot\cdot\cdot C}, \tag{31}$$

hence

$$\Lambda^C_{Ab} = -\tfrac{1}{4}\Gamma_{a\mathfrak{a}b}\,\alpha^{a\mathfrak{a}C}_{\cdot\cdot\cdot A} - 1/16\alpha^{\cdot D}_{\mathfrak{a}\cdot B}(\partial_b\,\alpha^{\cdot B}_{\mathfrak{e}\cdot D})\alpha^{a\mathfrak{e}C}_{\cdot\cdot\cdot A}. \tag{32}$$

If there is only one α given we can take the local unit vectors and spin-coordinates in such a way that the ordinary differentials of the components of α vanish. If there are more α, for instance one for one particle and the other for another particle, this is not possible for the components of both α's. In this case the last term to the right may get some physical interpretation. If the differentials vanish for any particular choice of unit vectors and

spin coordinates they will stay zero under: (1) Orthogonal transformations of unit vectors combined with the corresponding transformation of spin coordinates, (2) synchrone orthogonal transformation of unit vectors, (3) synchrone transformations of spin coordinates and (4) all transformations that can be built up from those already mentioned. Assuming the differentials to vanish we get

$$\Lambda^{C}_{Aj} = -\tfrac{1}{4}\Gamma_{acb}\,\alpha^{acC}_{\ldots A}. \tag{33}$$

To see the difference from the four-dimensional case we decompose with respect to the planes of α^4:

$$\begin{aligned}
\nabla_b\psi^c &= \partial_b\psi^c + \Lambda^c_{ab}\psi^a + \Lambda^c_{\mathfrak{A}b}\psi^{\mathfrak{A}} + \frac{ei}{ch}\varphi_b\psi^c,\\
\nabla_b\psi^{\mathfrak{C}} &= \partial_b\psi^{\mathfrak{C}} + \Lambda^{\mathfrak{C}}_{\mathfrak{A}b}\psi^{\mathfrak{A}} + \Lambda^{\mathfrak{C}}_{ab}\psi^a + \frac{ei}{ch}\varphi_b\psi^{\mathfrak{C}}.
\end{aligned} \tag{34}$$

Then from (33) we have

$$\begin{aligned}
\Lambda^c_{ab} &= -\tfrac{1}{4}\,\Gamma_{\mathfrak{A}ab}\,\beta^{\mathfrak{A}c}_{\ldots a},\\
\Lambda^{\mathfrak{C}}_{\mathfrak{A}b} &= -\tfrac{1}{4}\,\Gamma_{\mathfrak{A}ab}\,\overline{\beta}^{a\mathfrak{C}}_{\ldots\mathfrak{A}},\\
\Lambda^c_{\mathfrak{A}b} &= +\tfrac{1}{2}i\Gamma_{4ab}\,\beta^{\mathfrak{C}c}_{\ldots\mathfrak{A}},\\
\Lambda^{\mathfrak{C}}_{ab} &= -\tfrac{1}{2}i\Gamma_{4ab}\,\overline{\beta}^{\mathfrak{C}}_{\ldots a},
\end{aligned} \tag{35}$$

the terms with $\Lambda^c_{\mathfrak{A}b}$ and $\Lambda^{\mathfrak{C}}_{ab}$ in (34) are due to the fact *that the two planes of α^4 are no longer invariant under the displacement.* The Dirac-equations are

$$\begin{aligned}
\frac{h}{i}\overline{\beta}^{\mathfrak{C}}_{\ldots b}\left(\partial_j\psi^b + \Lambda^b_{aj}\psi^a + \frac{ei}{ch}\varphi_j\psi^b\right) + mc\psi^{\mathfrak{C}} &= 0,\\
\frac{h}{i}\beta^{\mathfrak{A}}_{\ldots\mathfrak{B}}\left(\partial_j\psi^{\mathfrak{B}} + \Lambda^{\mathfrak{B}}_{\mathfrak{A}j}\psi^{\mathfrak{A}} + \frac{ei}{ch}\varphi_j\psi^{\mathfrak{B}}\right) - mc\psi^c &= 0.
\end{aligned} \tag{36}$$

The equation

$$\frac{h}{i}\alpha^b\nabla_b\psi^C = 0 \tag{37}$$

can be decomposed into [6]

$$
\begin{aligned}
&\frac{h}{i}\bar{\beta}^{\mathfrak{A}}_{\ldots\mathfrak{b}}\left(\partial_j\varphi^{\mathfrak{b}}+\Lambda^{\mathfrak{b}}_{\mathfrak{a}j}\psi^{\mathfrak{a}}+\frac{ei}{ch}\varphi_j\psi^{\mathfrak{b}}\right)+h\nabla_4\psi^{\mathfrak{A}}\\
&\qquad+\tfrac{1}{2}h\Gamma_{4\mathfrak{k}\mathfrak{l}}\bar{\beta}^{\mathfrak{k}\mathfrak{l}\mathfrak{A}}_{\ldots\mathfrak{n}}\psi^{\mathfrak{n}}+\tfrac{1}{2}h\Gamma_{4\mathfrak{k}\mathfrak{l}}g^{\mathfrak{k}\mathfrak{l}}\psi^{\mathfrak{A}}=0,\\
&\frac{h}{i}\beta^{\mathfrak{x}}_{\ldots\mathfrak{B}}\left(\partial_j\psi^{\mathfrak{B}}+\Lambda^{\mathfrak{B}}_{\mathfrak{n}j}\psi^{\mathfrak{n}}+\frac{ei}{ch}\varphi_j\psi^{\mathfrak{B}}\right)-h\nabla_4\psi^{\mathfrak{x}}\\
&\qquad-\tfrac{1}{2}h\Gamma_{4\mathfrak{k}\mathfrak{l}}\beta^{\mathfrak{k}\mathfrak{l}\mathfrak{x}}_{\ldots\mathfrak{a}}\psi^{\mathfrak{a}}-\tfrac{1}{2}h\Gamma_{4\mathfrak{k}\mathfrak{l}}g^{\mathfrak{k}\mathfrak{l}}\psi^{\mathfrak{x}}=0.
\end{aligned}
\tag{38}
$$

In order that these equations be equivalent to (36) it is necessary and sufficient that:

$$
\begin{aligned}
&\alpha)\ \nabla_4\psi^{\mathfrak{A}}=0; \qquad \nabla_4\psi^{\mathfrak{x}}=0,\\
&\beta)\ \Gamma_{4[kl]}=0,\\
&\gamma)\ g^{ij}\Gamma_{4ij}=\frac{2mc}{h}.
\end{aligned}
\tag{39}
$$

The geometrical meaning of (39β) is that the congruence of $\underset{4}{e}^{\nu}$ is V_4-*normal*, in other words *time-space is holonomic in* V_5. The geometrical meaning of (39α) is that the spin-vector density ψ is *pseudo-parallel* along the curves of the congruence $\underset{4}{e}^{\nu}$. The planes E_2 and $\bar{E}_2$ are pseudo-parallel along these curves if and only if $\Lambda^{\mathfrak{A}}_{\mathfrak{a}4}$ and $\Lambda^{\mathfrak{x}}_{\mathfrak{B}4}$ vanish, that is if Γ_{4kl} vanishes. This means geometrically that the curves of $\underset{4}{e}^{\nu}$ are then *geodesics*. The Γ_{kl} are equal to the components of the so called second fundamental tensor of the V_4 of time-space

$$
h_{ij}\overset{\circ}{=}\Gamma_{kl} \tag{40}
$$

and $h'=g^{ij}h_{ij}$ is the invariant that vanishes only for minimal V_4. Hence (39γ) means that this invariant h' is constant and equal to $-\dfrac{2mc}{h}$.

Having introduced the V_5 to bring in the mass-term in a geometrical way it seems natural to suppose that the V_5 is

[6] The terms connected with the field being introduced by considering ψ as a pseudovector density as was pointed out in I.

euclidean. Then we have according to the theorem of Gauss

$$K_{ijkl} = -2h_{i[k}h_{l]j}. \tag{41}$$

and hence

$$K = g^{ik}K_{ijk}^{\ \ j} = -h^{i}_{\ j}h^{j}_{\ i} + \frac{4m^2c^2}{h^2}. \tag{42}$$

So we have proved now the theorem:

The Dirac-equations for Riemannian space-time with electromagnetic field and mass can be written in the form of equations without field or mass

$$\alpha^b \nabla_b \psi = 0 \tag{43}$$

in an R_5. Here α^b is the set of Dirac-numbers defined by (4) and ∇ is the symbol of a covariant differentiation whose parameters are given by (35). The congruence of e_4^r is V_4-normal, the spin-vector field ψ is pseudo-parallel along the curves of this congruence. The V_4 plays the role of time-space, the scalar h' of its second fundamental tensor h_{ij} is a constant and equal to $-\frac{2mc}{h}$.

There is a choice of h_{ij} that has a special interest, the choice that leads to a V_4 with constant curvature:

$$\Gamma_{ikj} = h_{ij} = -\frac{mc}{2h}g_{ij}. \tag{44}$$

In this case we have

$$K = -\frac{m^2c^2}{h^2} + \frac{4m^2c^2}{h^2} = 3\frac{m^2c^2}{h^2} \tag{45}$$

and the curvature of the V_4 is

$$\underset{0}{K} = -1/12K = -\frac{m^2c^2}{4h^2}. \tag{46}$$

The equation of a possible V_4 of this kind in the R_5 is

$$x^2 + y^2 + z^2 - c^2t^2 - u^2 = -\frac{4h^2}{m^2c^2}. \tag{47}$$

All V_4 are congruent, the curves of the congruence $\underset{4}{e^r}$ are no geodesics and the planes E_2 and $\bar{E}_2$ are not pseudo-parallel along these curves.

A NEW DEDUCTION OF THE GAUSSIAN DISTRIBUTION

By Norbert Wiener

It is well known that under fairly general conditions, an arbitrary law of distribution for a simple error will lead asymptotically to the Gaussian distribution for composite errors made up of large numbers of independent simple errors. An excellent treatment of this question is to be found in Paul Lévy's *Calcul des probabilités*. The present paper belongs to the domain of ideas there treated, and presents a related but perfectly distinct property of the Gaussian distribution. This property may be stated as follows: instead of compounding simple errors without number, let us see if we can resolve each apparently simple error into the composition of similar components with a perfectly arbitrary distribution. Under certain auxiliary assumptions of a very broad nature, we shall find, either that our process of resolution terminates at a stage where the mean square error is still finite, or the original distribution is precisely Gaussian. In more exact language, *let* $f(x, a)$ *be defined for* $-\infty < x < \infty$, $0 < a < A$, *let it be even and non-negative, and let*

$$\int_{-\infty}^{\infty} f(x, a)dx = 1. \tag{1}$$

Let

$$\int_{-\infty}^{\infty} x^2 f(x, a)dx \tag{2}$$

and

$$\int_{-\infty}^{\infty} [f(x, a)]^2 dx \tag{3}$$

be finite for each a. *Let*

$$\lim_{a \to 0} f(x, a) = 0, \qquad [x \neq 0] \tag{4}$$

and let

$$\int_{-\infty}^{\infty} f(x, a)f(y - x, b)dx = f(y, a + b) \tag{5}$$

for every positive a and b for which $a + b < A$. *Then*

$$f(x, a) = \frac{1}{\sqrt{2\pi Ha}} e^{-(x^2/4Ha)}, \tag{6}$$

where H is a constant.

To prove this, let us notice that, as Lévy has pointed out,

$$\begin{aligned}\int_{-\infty}^{\infty} y^2 f(y, a + b)dy &= \int_{-\infty}^{\infty} f(x, a)dx \int_{-\infty}^{\infty} y^2 f(y - x, b)dy \\ &= \int_{-\infty}^{\infty} f(x, a)dx \int_{-\infty}^{\infty} (z^2 + 2zx + x^2)f(z, b)dz \\ &= \int_{-\infty}^{\infty} x^2 f(x, a)dx \int_{-\infty}^{\infty} f(z, b)dz \\ &\qquad + \int_{-\infty}^{\infty} f(x, a)dx \int_{-\infty}^{\infty} z^2 f(z, b)dz \\ &= \int_{-\infty}^{\infty} x^2 f(x, a)dx + \int_{-\infty}^{\infty} x^2 f(x, b)dx. \end{aligned} \tag{7}$$

Again, whatever B may be,

$$\begin{aligned}\overline{\lim_{\epsilon \to 0}} \int_{-\infty}^{\infty} x^2 f(x, \epsilon)dx \leq \overline{\lim_{\epsilon \to 0}} \left[\int_{-\infty}^{-B} + \int_{B}^{\infty} \right] x^2 f(x, \epsilon)dx \\ + \overline{\lim_{\epsilon \to 0}} \int_{-B}^{B} x^2 f(x, \epsilon)dx. \end{aligned} \tag{8}$$

Let us notice that by (4),

$$\overline{\lim_{\epsilon \to 0}} \left[\int_{-\infty}^{-B} + \int_{B}^{\infty} \right] x^2 f(x, \epsilon)dx = 0, \tag{9}$$

and hence that

$$\overline{\lim_{\epsilon \to 0}} \int_{-\infty}^{\infty} x^2 f(x, \epsilon)dx \leq \overline{\lim_{\epsilon \to 0}} \int_{-B}^{B} x^2 f(x, \epsilon)dx. \tag{10}$$

Thus

$$\overline{\lim_{\epsilon \to 0}} \int_{-\infty}^{\infty} x^2 f(x, \epsilon)dx \leq B^2, \tag{11}$$

which is impossible for all B unless

$$\lim_{\epsilon \to 0} \int_{-\infty}^{\infty} x^2 f(x, \epsilon) dx = 0. \tag{12}$$

Combining this continuity condition with (7), we see that

$$\int_{-\infty}^{\infty} x^2 f(x, a) dx = Ha, \tag{13}$$

where H is some constant.

Let $G(x)$ be analytic and $O(x^{-2})$ at infinity. The expression

$$\left[\int_{-\epsilon}^{\epsilon} G(x) f(x, \epsilon) dx - G(o) \right] \Big/ \int_{-\epsilon}^{\epsilon} x^2 f(x, \epsilon) dx \tag{14}$$

will be a weighted average of the difference-quotients

$$\frac{G(x) - 2G(o) + G(-x)}{x^2} \tag{15}$$

for $x < C$, and will be bounded. Thus by (4),

$$\lim_{\epsilon \to 0} \left\{ \left[\int_{-\epsilon}^{\epsilon} G(x) f(x, \epsilon) dx - G(o) \right] \Big/ \int_{-\epsilon}^{\epsilon} x^2 f(x, \epsilon) dx \right.$$
$$\left. - \left[\int_{-\infty}^{\infty} G(x) f(x, \epsilon) dx - G(o) \right] \Big/ \int_{-\infty}^{\infty} x^2 f(x, \epsilon) dx \right\} = 0. \tag{16}$$

Thus the ultimate bounds of oscillation of

$$\left[\int_{-\infty}^{\infty} G(x) f(x, \epsilon) dx - G(o) \right] \Big/ \int_{-\infty}^{\infty} x^2 f(x, \epsilon) dx \tag{17}$$

will not be exterior to the range of (15) for values of x between $-C$ and C. This is however true for all C, however small, and yields $G''(0)$ for the limit of (17) as $\epsilon \to 0$. More generally,

$$\lim_{\epsilon \to 0} \left[\int_{-\infty}^{\infty} G(x + \xi) f(x, \epsilon) dx - G(\xi) \right] \Big/ \int_{-\infty}^{\infty} x^2 f(x, \epsilon) dx = G''(\xi). \tag{18}$$

Let us now assume further that $G'(x)$, $G''(x)$, and $G'''(x)$ exist for every x and belong to L_2. Let us put

$$\gamma(u) = \int_{-\infty}^{\infty} G(x)e^{-iux}dx \tag{19}$$

and

$$\varphi(u, a) = \int_{-\infty}^{\infty} f(x, a)e^{-iux}dx. \tag{20}$$

By condition (1), $\varphi(u, a)$ is continuous in u. By Plancherel's theorem,

$$G(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \gamma(u)e^{iux}du \tag{21}$$

and

$$f(x, a) = \underset{A\to\infty}{\text{l.i.m.}} \frac{1}{2\pi}\int_{-A}^{A} \varphi(u, a)e^{iux}du. \tag{22}$$

By the Parseval theorem for Fourier transforms,

$$\varphi(u, a)\varphi(u, b) = \varphi(u, a + b), \tag{23}$$

and by (1) and (4),

$$\lim_{\epsilon\to 0} \varphi(u, \epsilon) = 1. \tag{24}$$

Thus where

$$\varphi(u, a) \neq 0,$$

$$\varphi(u, a) = e^{a\psi(u)}, \tag{25}$$

where $\psi(u)$ is an even function. By a further use of the Parseval[1] theorem,

$$\frac{1}{\epsilon}\left[\int_{-\infty}^{\infty} G(x + \xi)f(x, \epsilon)dx - G(\xi)\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \gamma(u)e^{iu\xi}\left(\frac{e^{a\psi(u)} - 1}{2}\right)du. \tag{26}$$

However, by (18) and (13),

$$\lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\int_{-\infty}^{\infty} G(x + \xi)f(x, \epsilon)dx - G(\xi)\right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \gamma(u)e^{iu\xi}(-Hu^2)du. \tag{27}$$

Combining this with (26),

$$\lim_{\epsilon \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} \gamma(u) e^{iu\xi} \left[\frac{e^{\epsilon\psi(u)} - 1}{\epsilon} + Hu^2 \right] du = 0. \tag{28}$$

Now let u_0 be any value of u for which $\varphi(u, a) \neq 0$. Let (u_1, u_2) be a closed interval containing u_0, which we may choose so short that it contains no point for which: $\varphi(u, a) = 0$. Over this interval we shall uniformly find

$$\lim_{\epsilon \to 0} \frac{e^{\epsilon\psi(u)} - 1}{\epsilon} = \psi(u). \tag{29}$$

Hence if $\gamma(u)$ is so chosen, as is possible, not to vanish at u_0, to vanish outside (u_1, u_2), and to possess continuous derivatives at all points of orders up to and including the third, $G(x)$ will satisfy all the conditions which we have set for it, and

$$\lim_{\epsilon \to 0} \frac{1}{2\pi} \int_{-\infty}^{\infty} \gamma(u) e^{iu\xi} \left[\frac{e^{\epsilon\psi(u)} - 1}{\epsilon} - \psi(u) \right] du = 0. \tag{30}$$

Combining (28) and (30), we see that

$$\psi(u_0) = -Hu_0^2, \tag{31}$$

and since this holds for all values of u, except those of a set of zero measure,

$$\begin{aligned} f(x, a) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-Hau_0^2 - iu_0 x} du_0 \\ &= \frac{1}{\sqrt{2\pi Ha}} e^{-(x^2/4\pi Ha)}. \end{aligned} \tag{32}$$

It is perhaps worth while to comment that the whole argument turns on the positiveness of $f(x, \epsilon)$ for small values of ϵ. Thus the question of the positiveness of the Fourier transform, which is so difficult to treat in the case of a single function, offers a much easier approach in the case of the complete set of powers of a given function.

ALGEBRAIC MATRIC EQUATIONS[1]

By Philip Franklin

1. Introduction. We begin with equations of the type

$$P(X) = A,$$

where A is a given square matrix of order n, X is a similar matrix to be found, and $P(X)$ is a polynomial with scalar coefficients.[2] For the special case $X^n = A$, with A non-singular, the existence of a solution is well known,[3] and recently Roth[2] found all the solutions of the polynomial equation which are expressible as polynomials in A. In our treatment, we make use of a slight modification of the usual canonical forms for matrices, to find the solutions X expressible as polynomials in A. We also develop a systematic procedure for finding all the roots. When the characteristic equation of A has all its roots simple, the polynomial equation in X always has a finite number of roots. In other cases, however, there may be none, or an infinite number of roots. There are, however, only a finite number of roots which are not equivalent matrices.

We obtain another treatment of the equation $X^n = A$, when A is non-singular by generalizing the logarithmic function to matrices along lines suggested by Phillips.[4] This illustrates the extension of our methods from polynomials to integral functions.

We may generalize the polynomial matric equation by allowing the coefficients of $P(X)$ to be square matrices[5] expressible as

[1] Presented, in part, to the American Mathematical Society.

[2] Cf. W. E. Roth, "The Matric Equation $P(X) = A$," *Transactions of the American Mathematical Society*, vol. 30 (1928), p. 579 and the references there given.

[3] L. E. Dickson, *Modern Algebraic Theories*, Chicago, 1926, p. 121.

[4] H. B. Phillips, "Functions of Matrices," *American Journal of Mathematics*, vol. 41 (1919), p. 278.

[5] W. E. Roth, *Bulletin of the American Mathematical Society*, vol. 37 (1931), p. 670, Abstract no. 295.

polynomials or power series in a matrix A, or to be any rectangular matrices[6] of m rows and n columns. But, for any such matric equation, we may write a polynomial equation with scalar coefficients, which includes among its roots all of the roots of the original equation. Thus, theoretically, we may solve the equation with matric coefficients by selecting from the general solution of the scalar equation the proper subset of values. The method is only practical when we are seeking a restricted type of solution, as for example, solutions polynomials in, or commutative with, a given matrix.

We conclude with an application to generalized conjugate matrices.[7]

2. Canonical Forms. For our purposes, it is convenient to use a normal form for a matrix differing slightly from that ordinarily given. Preparatory to the introduction of this form, we recall some of the results of the theory of elementary divisors.[8] For any square matrix A, there exists a non-singular matrix Q such that

$$A = Q \cdot M \cdot Q^{-1}, \tag{2.1}$$

where M is any matrix with the same elementary divisors. In particular we may specialize M once for all to a canonical form, admitting all possible elementary divisors. One such form is a series of square blocks along the main diagonal, surrounded by zeros, with each block of the type

$$M_i = \begin{Vmatrix} a_i & 1 & 0 & \cdots & 0 \\ 0 & a_i & 1 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & a_i \end{Vmatrix}, \tag{2.2}$$

corresponding to an elementary divisor $(\lambda - a_i)^e$. In place of

[6] W. E. Roth, "On the Unilateral Equation in Matrices," *Transactions of the American Mathematical Society*, vol. 32 (1930), p. 61.

[7] "Generalized Conjugate Matrices," *Annals of Mathematics*, vol. 23 (1921), p. 97.

[8] M. Bocher, *Introduction to Higher Algebra*, New York, 1919, pp. 283–295.

these we shall admit blocks of the more general type:

$$N_i = \begin{Vmatrix} a_i & b_i & c_i & \cdots & t_i \\ 0 & a_i & b_i & \cdots & s_i \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdots & a_i \end{Vmatrix}, \qquad b_i \neq 0, \tag{2.3}$$

which evidently have the same elementary divisors as the blocks M_i. For any square matrix A, as before, we may find a non-singular matrix R, such that:

$$A = R \cdot N \cdot R^{-1}, \tag{2.4}$$

where N is built up of blocks N_i, surrounded by zeros. The a_i are fixed by the matrix A, the b_i are any numbers not zero, and the remaining constants, c_i etc., are absolutely arbitrary.

The advantage of the less definite, but more flexible, form (2.3) over the form (2.2) is at once apparent when we perform simple algebraic operations on our matrices. For, if two matrices are reduced by the same matrix R to normal forms built up of blocks N_i of the same order, their sum or product will be reduced to a form built up of blocks which are the sums and products of the corresponding N_i. Similarly if a matrix is multiplied by a scalar, each block of its normal form will be multiplied by a scalar. But, if the original blocks N_i were of the restricted form M_i, they would in general no longer be of this restricted form after the addition, multiplication, or multiplication by a scalar, and after a number of such operations would be of the general form N_i. On the other hand, the form N_i is substantially preserved under these operations, the only exception being that the result may be of the form given in (2.3) with $b_i = 0$. The product of a matrix N_i with $a, b, c \cdots$ as the elements of its first row by one with $x, y, z \cdots$ has as the elements of its first row the successive coefficients of

$$(a + bX + cX^2 + \cdots)(x + yX + zX^2 + \cdots),$$

i.e.:

$$ax, \quad ay + bx, \quad az + by + cx, \quad \cdots. \tag{2.5}$$

Consequently, the nth power of a matrix N_i with $a, b, c \cdots$ for the elements of the first row will be obtained from

$$(a + bX + cX^2 + \cdots)^n,$$

and will have as the elements of its first row:

$$(2.6)\quad a^n,\quad nba^{n-1},\quad \frac{n(n-1)}{2!}b^2a^{n-2} + nca^{n-1},$$

$$\frac{n(n-1)(n-2)}{3!}b^3a^{n-3} + \frac{n(n-1)}{1\cdot 1}bca^{n-2} + nda^{n-1},\quad \cdots.$$

In consequence of this, if $P(X)$ is a polynomial with scalar coefficients, the elements of the first row of the matrix $P(N_i)$, for the N_i just described, will be:

$$(2.7)\quad P(a),\quad bP'(a),\quad \frac{b^2}{2!}P''(a) + cP'(a),$$

$$\frac{b^3}{3!}P'''(a) + bcP''(a) + dP'(a),\quad \cdots.$$

This will be of type N_i, unless $P'(a)$ vanishes.

If A is reduced by (2.4), we shall evidently have

$$(2.8)\qquad P(A) = R\cdot P(N)\cdot R^{-1},$$

where $P(N)$ consists of a series of blocks $P(N_i)$, each found from (2.7).

3. Commutative Matrices. For later application, we shall now characterize all matrices which commute with a given matrix, which we reduce to the classical canonical form. Let P_1 be a matrix reduced to canonical form M_1, so that:

$$P_1 = R\cdot M_1\cdot R^{-1},$$

and P_2 be any other matrix commutative with it. Define a new matrix T by the relation:

$$T = R^{-1}\cdot P_2\cdot R.$$

Then

$$P_2 = R \cdot T \cdot R^{-1},$$

so that T is the matrix to which P_2 is reduced by R, and the assumed commutativity,

$$P_1 \cdot P_2 = P_2 \cdot P_1,$$

necessitates

$$M_1 \cdot T = T \cdot M_1,$$

so that T commutes with M_1, and the problem is reduced to finding all matrices which commute with one in normal form.

Next consider a matrix

$$\begin{Vmatrix} A & 0 \\ 0 & B \end{Vmatrix}$$

where A is a square matrix of n rows, while B is one of m rows, and the zeros stand for rectangular matrices with all their elements zero. Let

$$\begin{Vmatrix} P & Q \\ R & S \end{Vmatrix}$$

be a matrix of similar structure, so that P and S are square matrices of order n and m respectively, which commutes with it. We evidently must have:

(3.1)
$$\begin{aligned} A \cdot P &= P \cdot A, \\ A \cdot Q &= Q \cdot B, \\ B \cdot R &= R \cdot A, \\ B \cdot S &= S \cdot B. \end{aligned}$$

In particular, if A and B each have all the elements below the main diagonal zero, and each element of the main diagonal of A is distinct from every element of the main diagonal of B, these conditions make

$$Q = R = 0,$$

and the pairs A, P; Q, S commutative. We may apply this result to M_1, by grouping together all blocks M_i with equal

elements in the main diagonal into new blocks M_i'. For example, one such block might be:

$$M_1' = \begin{Vmatrix} a & 1 & 0 & 0 & 0 & 0 \\ 0 & a & 1 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 1 & 0 \\ 0 & 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 & 0 & a \end{Vmatrix}. \tag{3.2}$$

Then the result just found shows that any matrix T commutative with M_3 must consist of square blocks T_i of the same order as the corresponding M_i' along the main diagonal, surrounded by zeros. Moreover, each block T_i commutes with the corresponding block M_i', so that our problem is reduced to finding its solution for such blocks.

If the block M_i' under consideration consists of a single block of m rows, so that it is an M_i of the form given in (2.2), it is readily verified that the only matrix T_i commutative with it is a single block U_m of m rows similar in structure to the N_i of (2.3), except that the elements of its first row are arbitrary. If the block M_i' consists of two equal blocks each of m rows, then the fact just mentioned, together with the conditions (3.1) show that the only T_i commutative with it is one built up of four blocks U_m, similar to N_i. If the block M_i' consists of two unequal blocks, of m and n rows respectively, say $m > n$, a consideration of the conditions (3.1) shows that the corresponding T_i will have two blocks U_m and U_n along the main diagonal, surrounded by a rectangular block formed from the first n columns of a U_m above the main diagonal, and by one formed from the last n rows of a U_m below the main diagonal. In the general case M_i' is formed from a series of square blocks of $m_1, m_2, \cdots$ rows in order, and we may arrange them in decreasing order, so that

$$m_1 \geq m_2 \geq m_3 \cdots.$$

Then the corresponding T_i will have a series of blocks $U_{m_1}, U_{m_2}, \cdots$ along the main diagonal, surrounded above the main diagonal

by rectangular blocks consisting of the first m_i columns of a U_{m_i}, and below the main diagonal by rectangular blocks consisting of the last m_i rows of a U_{m_i}. The elements of all the blocks U mentioned are arbitrary, and independent of one another. As an illustration, the general matrix T_1 commutative with the matrix M_1' given in (3.2) will be of the form:

$$T_1 = \begin{Vmatrix} p_1 & q_1 & r_1 & p_{12} & q_{12} & p_{13} \\ 0 & p_1 & q_1 & 0 & p_{12} & 0 \\ 0 & 0 & p_1 & 0 & 0 & 0 \\ 0 & p_{21} & q_{21} & p_2 & q_2 & p_{23} \\ 0 & 0 & p_{21} & 0 & p_2 & 0 \\ 0 & 0 & p_{31} & 0 & p_{32} & p_3 \end{Vmatrix}. \tag{3.3}$$

It should be noted that, in general, two different matrices of this type will not commute.

When the M_i' are all identical with M_i, the result (2.7) shows that T_i will be expressible as a polynomial in M_i', and in fact, P_2 will be expressible as a polynomial in P_1. In general this will not be possible, and it will not be possible to find any third matrix such that both M_i' and T_i are polynomials in this third matrix. This was shown by Phillips,[9] who proved that the matrices:

$$\begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}, \qquad \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix}$$

are commutative, but are not expressible as polynomials in any third matrix. The discussion just given shows that this is the simplest example of such a pair. For an alternative proof, see example 7 at the end of section 5.

4. Solutions of $P(X) = A$, which are Polynomials in A. We are seeking solutions of the equation

$$P(X) = A, \tag{4.1}$$

where $P(X)$ is a polynomial with scalar coefficients, which are

[9] L.c.,[4] p. 275.

in turn expressible in the form:

$$X = L(A), \tag{4.2}$$

where L is a similar polynomial. If the matrix A is reduced to a canonical form N, built up of blocks N_i, by means of a matrix R, the result of (2.8) applied to (4.2) shows that R reduces X to blocks of similar type, except that some may have b_i zero. In this case these blocks could be further reduced, and a matrix R' would exist which, applied to X would split up these degenerate blocks into smaller ones. But, on applying the result of (2.8) to (4.1), we see that the matrix R' reduces A to smaller blocks than those in its canonical form, which is impossible. Thus, under the hypothesis of both the relations (4.1) and (4.2), there will be no degenerate blocks, and any matrix which reduces A to canonical form will also reduce X. Let, then R be such a matrix, so that:

$$A = R \cdot N \cdot R^{-1}, \tag{4.3}$$

$$X = R \cdot V \cdot R^{-1},$$

and, from (4.1):

$$P(V) = N.$$

If we call $a, b, c, \cdots$ the elements of a typical block N_i of N, and $x, y, z, \cdots$ those of a similar block V_i of V, the result (2.7) shows that:

$$P(x) = a, \tag{4.4}$$

$$yP'(x) = b,$$

$$\frac{y^2}{2!}P''(x) + zP'(x) = c,$$

etc.

For an N_i consisting of a single element, we have only the first of these equations, which may always be solved. For an N_i of more than one term, the second equation shows that we must have $P'(x)$ distinct from zero, since b is not zero, i.e. x must be a simple root of the first equation. This proves: [10]

[10] W. E. Roth, l.c.,[2] p. 588 gives the sufficiency of this condition. The necessity is stated in the footnote, p. 589, where a proof is outlined.

THEOREM I. *In order that the equation*

$$P(X) = A$$

have a solution expressible in the form

$$X = L(A),$$

it is necessary that the equation

$$P(x) = a_i$$

have at least one simple root for each root a_i yielding a block N_i of more than one term.

When an x is at hand satisfying the first of equations (4.4) for which $P'(x)$ is not zero, we may obtain y, z, u, etc. in succession from the equations (4.4) since at each stage the coefficient of the term to be found is distinct from zero. Thus we may find blocks V_i which combine to give a V satisfying $P(V) = N$, and hence from (4.3) an X satisfying (4.1). We must now show that, if we take the same root x for any two blocks V_i, corresponding to a pair of blocks N_i with identical a_i, this X will be a polynomial in A. Evidently (4.2) will be satisfied, provided a polynomial $L(A)$ is at hand for which

$$L(N) = V.$$

But this, by the result (2.7) is equivalent to:

$$\begin{aligned} L(a) &= x, \\ bL'(a) &= y, \\ \frac{b^2}{2!}L''(a) + cL'(a) &= z, \end{aligned} \tag{4.5}$$

etc. for each typical block. In the case for which all the a_i for different blocks N_i are distinct, these equations reduce to a set of conditions on the value of a polynomial $L(x)$ and its derivatives at certain points, and can always be satisfied by taking a polynomial of sufficiently high degree.

We must show that if two different blocks N_i have identical a_i, when we take the same simple root x for them, and then de-

termine the y, z, u etc. for each block from the equations (4.4), these values inserted in (4.5) do not lead to contradictory conditions. We first suppose we have two blocks of the same order, say N_1 and N_2, with identical a_i. By solving (4.4) we obtain two solutions, V_1 and V_2, with the same x, of the equations:

$$P(V_1) = N_1, \qquad P(V_2) = N_2.$$

Moreover, the matrices V_1 and V_2, with the same x, are equivalent, so that

$$V_2 = Q \cdot V_1 \cdot Q^{-1},$$

and the polynomial relations give in consequence of this:

$$N_2 = Q \cdot N_1 \cdot Q^{-1}.$$

If, now, $L(A)$ is any polynomial such that:

$$L(N_1) = V_1,$$

the two last equations show that we will also have:

$$L(N_2) = V_2.$$

This shows that the equations (4.5), which determine the values of the function and its derivatives uniquely at a_i, must give the same values for the two different blocks, N_1, N_2. The restriction that the blocks were of the same order is not an essential one, since a set of equations (4.4) or (4.5) m in number may always be considered as the first m of a set for a larger block, containing n equations, and the only part used in computing the first m derivatives.

This proves:[10]

THEOREM II. *The necessary condition of theorem I is also sufficient.*

Evidently for any a_i only appearing in a block N_i of one term, any root x may be used, otherwise we must use a simple root. This immediately leads to a method for getting all the solutions of the kind sought.[11]

[11] Cf. Roth, l.c.,[2] p. 591.

We note that while the b_i are always distinct from zero, some of the a_i may be zero, as in the:

Example 1.

$$X^2 + 2X = \begin{Vmatrix} 0 & 2 \\ 0 & 0 \end{Vmatrix} = A.$$

This has as solutions which are polynomials in A:

$$X = \begin{Vmatrix} 0 & 1 \\ 0 & 0 \end{Vmatrix} = \frac{A}{2}, \quad \text{or} \quad -\begin{Vmatrix} 2 & -1 \\ 0 & -2 \end{Vmatrix} = 2 - \frac{A}{2}.$$

5. General Solution of $P(X) = A$. We now proceed to develop an algorithm for finding all the solutions of the equation

$$P(X) = A.$$

It seems natural, at first sight, to consider the solutions which are commutative with A as a case intermediate between the polynomial and the general solution. However, since the coefficients of $P(X)$ are scalars, we have:

$$X \cdot P(X) = P(X) \cdot X,$$

and so all solutions of our equation necessarily commute with A.

If any solution X is reduced to canonical form V by means of R,

$$X = R \cdot V \cdot R^{-1}, \tag{5.1}$$

the matrix R will transform A to a matrix W,

$$A = R \cdot W \cdot R^{-1}, \tag{5.2}$$

where

$$P(V) = W.$$

By the result of (2.7), this is equivalent, for a typical block V_i, to

$$\begin{aligned} P(x) &= a, \\ yP'(x) &= b, \\ \frac{y^2}{2!}P''(x) + zP'(x) &= c, \end{aligned} \tag{5.3}$$

etc., where, as in (4.4), a, b, c, etc. are the elements of the first row of a typical block W_i, and x, y, z, etc., are the elements of the first row of the corresponding V_i. The blocks W_i are hence seen to be similar in structure to (2.3), except that we no longer have the restriction on the b, which here may be zero. We call the matrix built up of such blocks W_i a quasi-canonical form. We then see that if A is reduced to quasi-canonical form in all possible ways, and the equations (5.3) solved for each such reduction, all the solutions of the original equation will be found from (5.1).

If we know a single solution, X, of our original equation we may easily obtain from it all other solutions X_1, which are equivalent to it. For, if a square, non-singular matrix Q exists such that:

$$X_1 = Q \cdot X \cdot Q^{-1},$$

we deduce from

$$P(X) = P(X_1) = A,$$

that

$$A = Q \cdot A \cdot Q^{-1},$$

so that Q and A are commutative. To have X_1 a solution distinct from X, Q and X must not commute. Hence the solutions equivalent to X, and distinct from it are obtained by transforming it with all matrices which commute with A, but do not commute with X. These matrices Q are thus found from section 3. Evidently two distinct, equivalent, solutions are neither of them polynomials in A.

The structure of equations (5.3) shows that two quasi-canonical forms W consisting of blocks W_i with the corresponding blocks of the same size and mutually equivalent will lead to two sets of solutions for V, such that each member of the one set is equivalent to some member of the other set. Hence they may be found from one another by the method just described. Thus, in setting up the W_i from which the equations (5.3) are formed, we may confine our attention to the classical canonical form, together with those modifications of it leading to blocks of different sizes

with $b = 0$. For example, if a part of the canonical form were

$$\left\|\begin{matrix} a & 1 & 0 & 0 & 0 & 0 \\ 0 & a & 1 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 1 & 0 \\ 0 & 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 & 0 & a \end{matrix}\right\| \tag{5.4}$$

we would have to consider two forms in addition to this, and might take them as:

$$\left\|\begin{matrix} a & 1 & 0 & 0 & 0 & 0 \\ 0 & a & 1 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 0 & 1 \\ 0 & 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 & 0 & a \end{matrix}\right\| \text{ and } \left\|\begin{matrix} a & 0 & 1 & 0 & 0 & 0 \\ 0 & a & 0 & 1 & 0 & 0 \\ 0 & 0 & a & 0 & 1 & 0 \\ 0 & 0 & 0 & a & 0 & 0 \\ 0 & 0 & 0 & 0 & a & 0 \\ 0 & 0 & 0 & 0 & 0 & a \end{matrix}\right\| .$$

When all the a_i for the different blocks of the canonical form are distinct, there is, essentially only one reduction. In fact, in this case there are no solutions of the equation besides those expressible as polynomials in A already found in section 4.

When, in (5.3), $b \neq 0$, the solution proceeds exactly as in the case of (4.4). That is, for the first equation we take for x any root if the block consists of but one term. When the block has more than one term, we take any simple root to insure $P'(x)$ being different from zero. If there are no simple roots, there are no solutions for this particular reduction to the quasi-canonical form. After this root is selected, we determine y, z, u, etc., in turn, as before. If two different blocks have identical a_i, we are no longer obliged to take the same x for them, but may use any combination of simple roots for blocks of more than one term, and simple or multiple roots for blocks of a single term. When $b = 0$, we do not take $y = 0$, since this would lead to a V not in canonical form, and hence equivalent to the solution for a different choice of W_i. We here take a multiple root, which makes

$P'(x) = 0$. In general, this choice will lead to a contradiction when used in the later equations, but if it does lead to a system of equations solvable for y, z, u, etc.. we have a solution not equivalent to those already found.

We are now in a position to describe a systematic method of finding all the roots of the equation $P(X) = A$. We reduce A to the classical canonical form, and also to the quasi-canonical forms obtained from this by suitable interchanges of rows, and corresponding columns, as for (5.4). For each of these we solve the equations (5.3), if possible. We then obtain a complete set of non-equivalent solutions by using (5.1), for V from the canonical form, and the same R for V from the quasi-canonical form after the proper interchanges are made. Finally, we transform these by matrices Q commutative with A to obtain the solutions equivalent to these.

When the elementary divisors of A are all prime to one another, the only solutions are polynomial solutions. Hence the condition of Theorem I must be satisfied, and there are only a finite number of solutions.

When the elementary divisors are not all prime, there will still be only a finite number of non-equivalent solutions, but in general an infinite number of distinct solutions.

It is only in exceptional cases that the quasi-canonical forms yield any additional solutions.

We give some simple illustrations of the various cases.

Example 2. The equation

$$X^2 = \begin{Vmatrix} 0 & -1 \\ 0 & 0 \end{Vmatrix},$$

or the analogous equation

$$-X^2 + 2aX = \begin{Vmatrix} a & 1 \\ 0 & a \end{Vmatrix},$$

have no solutions.

Example 3. The equation

$$X^2 = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix},$$

has the polynomial solutions

$$\left\| \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix} \right\|, \qquad \left\| \begin{matrix} -1 & 0 \\ 0 & -1 \end{matrix} \right\|,$$

as well as the transforms of

$$\left\| \begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix} \right\|$$

by any matrix commutative with the right member, i.e. any non-singular matrix whatever. These may be reduced to the form:

$$\left\| \begin{matrix} a & (1-a^2)/b \\ b & -a \end{matrix} \right\| \text{ (including } \lim_{b=0} \text{ after } a^2 = 1).$$

Example 4. The equation

$$X^2 = \left\| \begin{matrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \right\|,$$

has the polynomial solutions

$$\pm \left\| \begin{matrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} \right\|$$

together with the transforms of

$$\pm \left\| \begin{matrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{matrix} \right\|$$

by matrices of the form

$$\left\| \begin{matrix} p_1 & q_1 & p_{12} \\ 0 & p_1 & 0 \\ 0 & p_{21} & p_2 \end{matrix} \right\|,$$

which may be reduced to:

$$\pm \left\| \begin{matrix} 1 & 1+ab & a \\ 0 & 1 & 0 \\ 0 & -2b & -1 \end{matrix} \right\|.$$

Example 5. The equation

$$X^2 = \begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix},$$

has the polynomial solution

$$\begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix},$$

together with the transforms of

$$\begin{Vmatrix} 0 & p \\ 0 & 0 \end{Vmatrix},$$

by any matrix, which may be reduced to the form:

$$\begin{Vmatrix} a & a^2/b \\ -b & -a \end{Vmatrix}.$$

Example 6. The equation

$$X^2 = \begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix},$$

has no polynomial solutions (cf. ex. 2.), but has the solutions:

$$\pm \begin{Vmatrix} 0 & c & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{Vmatrix},$$

arising from the quasi-canonical form, together with its transforms by the matrices used in example 4, which may be reduced to the form:

$$\begin{Vmatrix} 0 & a & b \\ 0 & 0 & 0 \\ 0 & 1/b & 0 \end{Vmatrix}.$$

Example 7. Let us attempt to find all matrices X, such that the matrix:

$$A = \begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix},$$

is expressible as a polynomial in X. Since X will be of the third order, it is a root of a cubic equation, i.e. its characteristic equation, and this may be used to reduce the polynomial which gives A to one of the second order. If it has equal roots, we may write:

$$\frac{X^3 - 2rX + r^2}{t^2} = A.$$

The roots of this (cf. example 6) are found to be:

$$X_1 = \begin{Vmatrix} r & a & tb \\ 0 & r & 0 \\ 0 & t/b & r \end{Vmatrix}.$$

If the quadratic has distinct roots, we may write:

$$\frac{X^2 - (r+s)X + rs}{t(r-s)} = A.$$

The roots of this are found to be:

$$X_2 = \begin{Vmatrix} r & t & 0 \\ 0 & r & 0 \\ 0 & 0 & r \end{Vmatrix}, \quad X_3 = \begin{Vmatrix} r & t - \dfrac{ab}{r-s} & a \\ 0 & r & 0 \\ 0 & b & s \end{Vmatrix},$$

together with the matrices obtained from them by interchanging r and s, and replacing t by $-t$. Obviously, if A is a first degree polynomial in a matrix X, this must have the form X_2. Thus, the only matrices X such that A is a polynomial in X are those of the form X_1, X_2, X_3.

If, in place of A above, we consider the matrix:

$$B = \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{Vmatrix},$$

obtained as a transform from A by interchanging the last two rows and last two columns, the matrices Y such that B is a polynomial in Y will be those obtained from X_1, X_2, X_3 by

interchanging the last two rows and last two columns. Since no matrix so obtained as a Y is a possible X, unless $t = 0$, there is no matrix such that A and B are both expressible as polynomials in it. This gives an alternative proof of the fact cited [9] at the end of section 3. The example also indicates a systematic procedure for finding all the matrices in terms of which a given matrix is a polynomial.

6. The Exponential and Logarithmic Function of a Matrix. When the general theory of analytic functions of matrices and their Taylor development [4] about a constant matrix is applied to the expansion of the exponential function about the zero matrix, we are led to a consideration of the series,

$$e^A = 1 + A + \frac{A^2}{2!} + \cdots, \tag{6.1}$$

the symbol on the left side of this equation being here merely the abbreviation for the series. If A is a square matrix of the mth order, and has all its terms numerically less than M, A^2 will have all its terms bounded by mM^2, and the terms of A^n will be bounded by $m^{n-1}M^n$, so that all the series for separate terms of the matrix will converge, and the series defines a new matrix for any matrix A.

If A and B are two matrices which *commute* with one another, the characteristic property of the exponential function is satisfied, namely:

$$e^A \cdot e^B = e^B \cdot e^A = e^{A+B} \quad \text{when} \quad A \cdot B = B \cdot A$$

If A and B do not commute, in general none of the equations just written will be satisfied.

If the matrix R transforms A to canonical form, N, as in (2.4), we will have:

$$e^A = R \cdot e^N R^{-1}.$$

The matrix e^N will be formed of blocks along the main diagonal, which may be evaluated as in (2.7). The elements of the first

row of a typical block will be:

$$(6.2) \qquad e^a, \quad be^a, \quad \frac{b^2}{2!}e^a + ce^a, \quad \frac{b^3}{3!}e^a + bce^a + de^a,$$

etc. Since $b \neq 0$ for a block of the canonical form N, and $e^a \neq 0$, the matrix e^N will always be in canonical form, and have the same size blocks as N.

If we define the logarithmic matrix function, as any inverse of the exponential function, then

$$X = \log A, \qquad \text{whenever} \qquad e^X = A.$$

Since any matrix X may be normalized by a matrix R to a normal form V, and R will then reduce A to a normal form N, we reduce A to such a normal form and then have to solve

$$e^V = N.$$

By the relations (6.2), this leads for a characteristic block with a, b, c, etc., as the elements of the first row for N, and x, y, z as those for V, to the equations:

$$(6.3) \qquad \begin{aligned} e^x &= a, \\ ye^x &= b, \\ \frac{y^2}{2!}e^x + ze^x &= c, \end{aligned}$$

etc. If $a \neq 0$, these equations may be solved for x, y, z, etc., in turn, while if $a = 0$, they have no solution. Thus a logarithmic function exists for any non-singular matrix A. For each block we take a particular branch of the ordinary scalar function $x = \log a$. Since, by (6.2), no quasi-normal forms need be considered, a single reduction of the matrix A to normal form will yield all the non-equivalent solutions. The argument used in section 5 shows that the solutions equivalent to those already found may be obtained by transforming them by matrices which commute with A, but do not commute with the solutions. As before, unless there are several blocks with equal a's, there will be no such transforms.

The logarithmic function just found is not analytic, but the solution includes those which are analytic, i.e. expansible in series about a suitable scalar, or matrix. These analytic solutions will be characterized by relations between the branches taken for the different blocks.

By means of the logarithmic function here defined we may find the general solution of the equation $X^n = A$, when A is any non-singular matrix. The method does not apply when A is singular, although we have seen (cf. examples 5, 6) that the equation may well have roots when A is singular.

The procedure we have employed to find the inverse of the exponential function may be extended to other integral functions. In general, we will not only have to solve a scalar equation for each characteristic root, but may have to find simple roots, as we did for polynomials in Theorem I. As an illustration, consider the equation:

$$\sin X = X - \frac{X^3}{3!} + \frac{X^5}{5!} + \cdots = \begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}.$$

This equation has no solution, since the scalar equation $\sin x = 1$ has no simple root. The result is confirmed by the fact that the determination of $\cos X$ from this value of $\sin X$ leads us, essentially, to the unsolvable equation of example 2.

7. The Generalized Characteristic Equation. The following theorem on matrices was proved by Phillips: [12]

If two sets of square matrices, A_s, $B_s (s = 1, 2, \cdots k)$, with elements a_{ij}^s and $b_{ij}^s (i, j = 1, 2, \cdots n)$ respectively, satisfy the equation

$$\sum_{s=1}^{k} A_s B_s = 0,$$

then, if all the A_s commute among themselves, they satisfy the equation:

$$\left| \sum_{s=1}^{k} A_s b_{ij}^s \right| = 0,$$

[12] L.c.,[4] p. 267.

and if all the B_s commute among themselves, they satisfy the equation:

$$|\sum_{s=1}^{k} B_s a_{ij}^{s}| = 0.$$

This theorem [12] leads to the well known fact that every square matrix satisfies its characteristic equation, when it is applied to the identity

$$A \cdot I - I \cdot A = 0,$$

where A is any square matrix, and I is the identity matrix of the same order.

Here we shall use it to obtain an equation with scalar coefficients from any polynomial (or power series) matric equation. For, let X and the $(m+1)$ coefficients A_s be square matrices of order n, and suppose that the equation

$$(7.1) \qquad A_m X^m + A_{m-1} X^{m-1} + \cdots + A_1 X + A_0 = 0,$$

is satisfied. Then, since the identity matrix and the powers of X all commute with one another, the theorem quoted shows that X satisfies [13] the equation:

$$(7.2) \qquad |a_{ij}^{m} X^m + a_{ij}^{m-1} X^{m-1} + \cdots a_{ij}^{1} X + a_{ij}^{0}| = 0.$$

We observe that if we had multiplied X and its powers on the left, to form the equation:

$$(7.3) \qquad X^m A_m + X^{m-1} A_{m-1} + \cdots X A_1 + A_0 = 0,$$

we would have found the same equation (7.2) from the first part of the theorem.

If the equation (7.2) which is usually of the mnth degree is not identically zero, we may find its complete solution by the algorithm developed in section 5. The solution will in general contain parameters, since the matrix N of section 5 is here a scalar times the identity, and so has all its characteristic roots equal. However, there will only be a finite number of possible

[13] The fact that (7.2) is a consequence of (7.1), and is an aid to the solution of the latter equation was announced by Professor C. C. MacDuffee at the meeting of the American Mathematical Society, September, 1931.

non-equivalent solutions, since the characteristic roots must be selected from the roots of the scalar equation corresponding to (7.2).

The solutions of (7.1) will be a subset of the solutions of (7.2), and accordingly we have merely to test these last in (7.1). Except for the restriction on the characteristic values, this last process may prove nearly as difficult as the testing of a general square matrix. However, if we are seeking solutions of the equation (7.1) which are commutative with a given matrix, or polynomials in a given matrix, we need only find solutions of equation (7.2) having these properties, and the problem of testing them in equation (7.1) becomes one practically capable of solution.

Instead of restricting the coefficients of the equation (7.1), or (7.3) to be square matrices, we may take them as rectangular matrices of n columns and t rows (n rows and t columns for (7.3)). This case is in one sense a special case of the problem for square coefficients, since if t is greater than n, the equation will be equivalent to a number of matrix equations with square coefficients. For example, we may form square coefficients from the first n, the second n, etc. as well as the last n rows (which will of course in general have some rows in common with the preceding equation). The roots of the original equation will be the common roots, if any, of this series of equations. In particular, if the corresponding equations (7.2) have no common roots, we know at once that the given equation has no solutions. If t is less than n, we form an equation with square coefficients by adding to each coefficient $n - t$ rows with arbitrarily chosen elements. The roots of any such modified equation will clearly satisfy the original equation.

In some cases, even when the determinant in the left member of equation (7.2) vanishes identically, we may still find such a scalar equation. Namely, when the vanishing of this determinant is due to the fact that one of the rows of *all* the coefficients is linearly dependent on the remaining rows. For, we have then

merely to replace the dependent row by arbitrary elements of an independent row, and proceed as before. However, the scalar equation may give a left member which vanishes identically without all the determinants of the coefficients being zero, which precludes the existence of any dependent rows. As an illustration of the first situation, we may write:

$$\begin{Vmatrix} 1 & 1 \\ 3 & 3 \end{Vmatrix} X - \begin{Vmatrix} 1 & 1 \\ 3 & 3 \end{Vmatrix} = 0,$$

which we replace by

$$\begin{Vmatrix} 1 & 1 \\ 0 & 1 \end{Vmatrix} X - \begin{Vmatrix} 1 & 1 \\ a & b \end{Vmatrix} = 0,$$

where a and b are arbitrary numbers. This has a scalar equation which does not reduce to an identity. The original equation has the solution

$$X = \begin{Vmatrix} 1-a & 1-b \\ a & b \end{Vmatrix}.$$

On the other hand, the equation:

$$\begin{Vmatrix} 1 & 0 \\ 0 & 0 \end{Vmatrix} X^2 + \begin{Vmatrix} 0 & 2 \\ 3 & 0 \end{Vmatrix} X + \begin{Vmatrix} 0 & 0 \\ 0 & 6 \end{Vmatrix} = 0,$$

has a scalar equation with vanishing left member. However, its solution

$$X = \begin{Vmatrix} 0 & -2 \\ a & b \end{Vmatrix},$$

has arbitrary, in general distinct, characteristic roots, which explains why no scalar equation limiting the values of these roots can be found. In this illustration A_0 and A_∞ both have zero determinants. This is a necessary condition for the determinant in the scalar equation to vanish identically, since the two determinants are the coefficients of the constant term, and of X^n in (7.2). This proves [14] that *the equation* (7.1) *or* (7.3) *can have solutions with variable characteristic roots only when the rank of* A_0 *and that of* A_∞ *are each less than* n.

[14] Cf.[8] l.c., p. 77.

8. Solutions of $P(A, X) = 0$ which are Polynomials in A. When we are dealing with an equation of the form:

$$\sum_{s=0}^{m} F_s(A)X^s = P(A, X) = 0, \tag{8.1}$$

where the coefficients are polynomials with scalar coefficients (or power series) in a matrix A, and we only wish solutions which are polynomials in A, we may use the procedure of section 4. Any matrix R which reduces A to normal form N, will reduce X to quasi-normal form V, and we have to solve the equation

$$P(N, V) = 0.$$

Each coefficient $F_s(N)$ will be in quasi-normal form. We call the elements of the first row of a typical block of N, a, b, c, etc. and of $F_s(N)$, a_s, b_s, c_s, etc., and the elements of the corresponding block of V, x, y, z, etc. Then for each block we have a set of equations analogous to (4.4):

$$\sum_{s=0}^{m} a_s x^s = \sum_{s=0}^{m} F_s(a)x^s = P(a, x) = 0, \tag{8.2}$$

$$\sum_{s=0}^{m} (a_s s y x^{s-1} + b_s x^s) = yP_x(a, x) + bP_a(a, x) = 0,$$

$$\frac{y^2}{2!}P_{xx} + zP_x + byP_{ax} + cP_a + \frac{b^2}{2!}P_{aa} = 0,$$

etc. In these equations, subscripts denote partial derivatives, and use has been made of (2.5), (2.6) and (2.7). Since y, z, etc., appear in these equations for the first time with the coefficient P_x, if this is not zero for some x which satisfies the first equation, we may solve for them in turn. This establishes:

THEOREM III. *A sufficient condition for the equation*

$$\sum_{s=0}^{m} F_s(A)X^s = P(A, X) = 0$$

to have a solution which is a polynomial in A is the existence of a simple root of the equation:

$$P(a, x) = 0,$$

for each root a which appears in a block of more than one term in N, the canonical form of A.

This condition is not necessary, as is seen from the equation:

$$X^2 = A^2,$$

where

$$A = \begin{Vmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{Vmatrix}.$$

For this equation is that of example 6, so that it has the solutions:

$$\begin{Vmatrix} 0 & a & b \\ 0 & 0 & 0 \\ 0 & 1/b & 0 \end{Vmatrix},$$

and when $b = \pm 1$, this reduces to

$$aA^2 \pm A,$$

which is a solution of the equation, expressible as a polynomial in A.

If we reduce the matrix A to normal form, and in addition consider all transpositions of rows, and corresponding columns, which give quasi-normal forms, by solving in each case for solutions which are expressible as polynomials in A, we will have all such polynomial solutions of the original equation.

9. Generalized Conjugate Matrices. Many proofs in the theory of matrices may be simply carried out by introducing the concept of conjugate matrices.[15] For a given matrix, M_1, a set of $(n-1)$ matrices $M_2, M_3, \cdots M_n$ are called a set of *conjugate matrices* of M_1, provided, 1. the n matrices M_i all commute with one another, 2. the symmetric functions of these n matrices are all scalar matrices, and equal to the corresponding symmetric functions of the n scalar roots of the characteristic equation of M_1,

[15] Taber, O., "On Certain Identities in the Theory of Matrices," *American Journal of Mathematics*, vol. 13 (1891), p. 159. Bennett, A. A., "Some Algebraic Analogies in Matric Theory," *Annals of Mathematics*, vol. 23 (1921), p. 91.

each multiplied by the identity matrix of order n, and 3. each matrix M_i has the same characteristic equation as M_1.

It is not always possible to find a set of conjugates for a given matrix, but a set of *generalized conjugate* matrices, satisfying conditions 1. and 2. only, may always be found.[7] Since any set of conjugate matrices is also a set of generalized conjugate matrices, if we find all possible sets of generalized conjugate matrices, we will have included the conjugate sets, when they exist, and hence have a method of determining this existence question.

From their definition, the product of the n factors $X - M_i$ will reproduce the characteristic equation of M_1. Conversely, if a factorization of this type exists, with the M_i commutative, we have a generalized conjugate set. Hence all the generalized conjugate matrices are to be found among the roots of this equation, which commute with M_1. These may be found by the methods of section 5. For an equation with only a finite number of roots, the matching up is easily accomplished. However, we may do this systematically by noting that, if X and A are commutative matrices, the factorization of $X^n - A^n$ is analogous to that for the scalar function, and from this it follows that if the matric polynomial $P(A, X)$ vanishes for $X = M_1$, commutative with A, it is divisible by $X - M_1$. Hence to find a generalized conjugate set, we divide the characteristic equation by $X - M_1$, then select any root of this, M_2, commutative with M_1, and divide by $X - M_2$, etc. If at any stage we meet an equation with no roots commutative with the M_i already found, these will not form part of a generalized conjugate set. The number of canonical forms for the roots is finite at each stage, so that there are at most a finite number of non-equivalent sets M_i. The transform of any set, by a matrix commutative with M_1 but not with all the members of the set, will give a distinct set.

THE PROBLEM OF PLATEAU FOR TWO CONTOURS[1]

By Jesse Douglas

Contents

	PAGE
1. Introduction	315
2. General Formulas	322
3. Review of the One-contour Case	328
4. Formulas for the Two-Contour Case	334
5. The Functional A in Terms of Fourier Coefficients	337
6. The Minimizing Representation	340
7. Vanishing of the First Variation of $A(g_1, g_2; q)$	343
8. Proof of the Fundamental Identity	345
9. Interlacing Contours	349
10. Conformal Mapping	352
11. Arbitrary Jordan Contours	354

1. Introduction

The problem of Plateau is to prove the existence of a minimal surface with a given boundary. The early work on this problem, by Riemann, Weierstrass and Schwarz, was confined in the main to a polygonal contour and was carried through only for special skew polygons (e.g., the skew quadrilateral obtained by removing a pair of opposite edges from a regular tetrahedron, Schwarz). The work of S. Bernstein, Ch. H. Müntz and A. Haar considered the minimal surface in the form $z = f(x, y)$ and assumed the contour to have a simply-covered convex projection on the xy-plane.

It is only in the last few years that the Plateau problem has been treated for a contour of general shape. The work in this

[1] Presented to the American Mathematical Society: "A General Formulation of the Problem of Plateau," Oct. 26, 1929; "Solution of the Problem of Plateau for Two Contours," Sept. 27, 1930; abstracts in *Bulletin American Mathematical Society*, vol. 36, p. 50, p. 797.

The seven main theorems established in this paper have been communicated to the National Academy of Sciences and will be published in its *Proceedings*, vol. 18, no. 1 (Jan. 1932).

direction was by R. Garnier, the present writer [2] and T. Radó. The problem was first definitively solved in the work of the present writer; not only did this establish the most general result: that every continuous, closed, non-self-intersecting curve (Jordan curve) in euclidean space of any number of dimensions is the boundary of a simply-connected portion of a minimal surface, but the method there developed—based on a certain simple functional $A(g)$ whose argument is a variable parametric representation of the given contour—furnishes, beyond a doubt, the key to the problem.

It may perhaps be regarded as only an incident that the present writer's work was the first to show that the number of dimensions of the containing space is indifferent; all the other work, without exception, is restricted to a three-dimensional containing space. But to use n dimensions cannot be considered entirely a banal generalization; for instance, it allows us to put $n = 2$, whereupon the conformal mapping of plane regions is included as a special case; this cannot be done with any of the other theories.

Up to now, the only work for two contours has been the classic case of two coaxial circles determining a catenoid, a treatment by Riemann [3] of two circles in parallel planes but not coaxial, and of any two polygons in parallel planes, and the treatment by Schwarz [4] (a special case of the preceding, more elaborately considered) of two regular polygons in the position of opposite bases of a regular prism. A general theory is therefore completely outstanding. The present paper, the first of any generality to

[2] *Bulletin American Mathematical Society*, vol. 33 (1927), pp. 143, 259; vol. 34 (1928), p. 405; vol. 35 (1929), p. 292; vol. 36 (1930), pp. 49–50, 189–190. A comprehensive and didactic presentation was published in *Transactions American Mathematical Society*, vol. 33, no. 1 (Jan. 1931): "Solution of the Problem of Plateau." This paper will be cited hereafter as (I). References to the other work on the problem may be found there.

The writer's work was also made abundantly known during the year 1929 by his lectures in mathematical seminars at Paris, Göttingen and Hamburg.

[3] Werke, Leipzig, 1892, pp. 329–333, 446–454.

[4] *Gesammelte Abhandlungen*, vol. 1, pp. 270–316.

deal with two contours, supplies this lack, extending to two contours the methods developed in the paper (I) for the case of a single contour. The main results here established may be described as follows.

Let Γ_1, Γ_2 denote the given contours, which we suppose to be Jordan curves in euclidean space of n dimensions. Even when explicit mention of the fact is omitted, *we assume throughout that Γ_1 and Γ_2 do not intersect one another.*

We consider the totality of ways of representing Γ_1, Γ_2 parametrically:

$$x_i = g_{1i}(\theta), \qquad x_i = g_{2i}(\theta) \qquad (i = 1, 2, \cdots, n),$$

θ denoting respectively the polar angle on two concentric circles C_1, C_2, with ratio of radii equal to q, $0 < q < 1$.

Then we define certain functionals:

$$A(g_1), \quad A(g_2), \quad A(g_1, g_2; q)$$

as follows. Let

$$(1.1) \qquad x_i = H_{1i}(u, v), \qquad x_i = H_{2i}(u, v), \qquad x_i = H_{12i}(u, v)$$

denote the harmonic functions determined respectively in the interior of C_1, the interior of C_2, the ring between C_1 and C_2 by the boundary values $g_{1i}(\theta)$, $g_{2i}(\theta)$, $g_{1i}(\theta)$ and $g_{2i}(\theta)$. These define harmonic (therefore analytic) surfaces bounded respectively by Γ_1, Γ_2, Γ_1 and Γ_2; let the fundamental quantities of these surfaces be denoted by E, F, G with the appropriate subscripts. Then by definition:

$$(1.2) \qquad \left.\begin{aligned} A(g_1) &= \tfrac{1}{2}\int\int (E_1 + G_1)dudv \\ A(g_2) &= \tfrac{1}{2}\int\int (E_2 + G_2)dudv \\ A(g_1, g_2; q) &= \tfrac{1}{2}\int\int (E_{12} + G_{12})dudv \end{aligned}\right\}$$
$$= \tfrac{1}{2}\int\int \sum_{i=1}^{n}\left[\left(\frac{\partial x_i}{\partial u}\right)^2 + \left(\frac{\partial x_i}{\partial v}\right)^2\right]dudv,$$

each double integral being taken over the appropriate region.

Since the integrand is an essentially positive quantity, the double integrals exist either as finite positive values or as $+\infty$.

When all possible parametric representations of Γ_1, Γ_2 and values of q are considered, the functionals A therefore have determinate lower bounds:

$$m(\Gamma_1), \quad m(\Gamma_2), \quad m(\Gamma_1, \Gamma_2)$$

$\leqq +\infty$ and $\geqq 0$ (in fact > 0, established after proof that these bounds are attained).

The quantities $m(\Gamma)$ are the same as those introduced by Lebesgue[5] on the basis of his definition of area (inferior limit of the areas of approaching polyhedral surfaces) and called by him the "minimum area" of the contour. It is however their above interpretation, based on the functional A, which will concern us *exclusively* in this paper. Lebesgue's definition is: $m(\Gamma_1)$ is the least limit which can be approached by the area of a simply-connected polyhedral surface whose polygonal boundary tends to Γ_1, with a similar definition, mutatis mutandis, for $m(\Gamma_1, \Gamma_2)$. These are also the lower bounds of the areas of continuous surfaces bounded respectively by Γ_1, by Γ_1 and Γ_2.[6] The main body of this paper is concerned with the case where the quantities $m(\Gamma)$ are finite; a sufficient condition for this is that the contours be rectifiable. Later (§ 11), extension is made by means of a limit process to a pair of arbitrary Jordan contours.

Remark. In order that all three quantities $m(\Gamma)$ be finite it is necessary and sufficient that $m(\Gamma_1, \Gamma_2)$ be finite. This is easily seen in § 11 in connection with the formulas (11.2).

With our definitions:

$$\begin{aligned} m(\Gamma_1) &= \text{lower bound } A(g_1), \\ m(\Gamma_2) &= \text{lower bound } A(g_2), \\ m(\Gamma_1, \Gamma_2) &= \text{lower bound } A(g_1, g_2; q), \end{aligned} \tag{1.3}$$

we have always

$$m(\Gamma_1, \Gamma_2) \leqq m(\Gamma_1) + m(\Gamma_2). \tag{1.4}$$

[5] "Intégrale, longueur, aire," *Annali di Matematica*, (3), vol. 7 (1902), p. 304.

[6] Even when explicit mention of the fact is omitted, we understand every surface bounded by Γ_1 or Γ_2 separately to be simply-connected, by Γ_1, Γ_2 jointly, doubly-connected.

This follows immediately from (5.6) or (6.3), according to which, for $q \to 0$,

$$A(g_1, g_2; q) \to A(g_1) + A(g_2).$$

With the Lebesgue interpretation of $m(\Gamma)$ the relation (1.4) is obvious. For if we connect by a straight line segment two interior vertices belonging respectively to two simply-connected polyhedral surfaces with boundaries tending to Γ_1 and Γ_2 and areas to $m(\Gamma_1)$ and $m(\Gamma_2)$, we may say that we have a doubly-connected polyhedral surface with area tending to $m(\Gamma_1) + m(\Gamma_2)$. Or again: if a point on a surface bounded by Γ_1 and with area arbitrarily close to $m(\Gamma_1)$ be connected by a straight line segment to a point on a surface bounded by Γ_2 and with area arbitrarily close to $m(\Gamma_2)$, the result may be considered a doubly-connected surface bounded by Γ_1, Γ_2, and its area is arbitrarily close to $m(\Gamma_1) + m(\Gamma_2)$.

Now suppose—with $m(\Gamma_1, \Gamma_2)$ finite (and hence, by the above remark, also $m(\Gamma_1)$ and $m(\Gamma_2)$ finite)—we have the strict inequality:

$$m(\Gamma_1, \Gamma_2) < m(\Gamma_1) + m(\Gamma_2), \tag{1.5}$$

then the main theorem of this paper is the following.

THEOREM I. *In euclidean space of n dimensions, let Γ_1, Γ_2 be two Jordan curves, not intersecting one another, with finite $m(\Gamma_1, \Gamma_2)$ and obeying the condition (1.5).*

Then there exists a pair of concentric circles C_1, C_2 and n functions $F_i(w)$ of the complex variable w, holomorphic in the ring between C_1 and C_2, obeying there identically the condition

$$\sum_{i=1}^{n} F_i'^2(w) = 0, \tag{1.6}$$

and whose real parts

$$x_i = \Re F_i(w) \tag{1.7}$$

attach continuously to boundary values on C_1, C_2:

$$x_i = g_{1i}(\theta), \qquad x_i = g_{2i}(\theta)$$

which are parametric representations of Γ_1, Γ_2.

(1.7) and (1.6) are the Weierstrass formulas for a minimal surface; in other words, this theorem asserts the existence of a doubly-connected minimal surface M bounded by Γ_1, Γ_2, or the solution of the problem of Plateau for those two contours.

The area of M is immediately seen to be $m(\Gamma_1, \Gamma_2)$, for the condition (1.6) is equivalent to

$$E = G, \qquad F = 0,$$

and then the functional A

$$\int\int \tfrac{1}{2}(E + G)dudv$$

is equal to the area

$$\int\int \sqrt{EG - F^2}\, dudv.$$

But to be able to infer that therefore M has the least area of any surface bounded by Γ_1, Γ_2, we need to know the identity of the Lebesgue interpretation of $m(\Gamma_1, \Gamma_2)$ with the definition (1.3) of the present paper. This could be established quite briefly, after the pattern of Part V of our first paper, with the use of the conformal mapping of a polyhedral surface, after Koebe. However, it is our intention to avoid Koebe's theory of conformal mapping of abstract Riemann manifolds,[7] and we therefore postpone further consideration of the least area property of the minimal surface whose existence is here established to a future paper, where we expect to give this part of the problem a direct treatment.

One reason for avoiding the use of any extraneous theory of conformal mapping in the present paper is that our methods enable us to give an independent foundation for this theory in

[7] In my paper "The Mapping Theorem of Koebe and the Problem of Plateau," this JOURNAL, vol. X (1931), pp. 106–130, it is shown that the polyhedral mapping theorem—part of the uniformization theory of Koebe—contains all that is difficult in the problem of Plateau; the establishing of that theorem practically *is* the solution of the problem of Plateau, it being only necessary to adjoin five lemmas, easy of statement and proof, to control the limit process from the polyhedrons to the minimal surface. An idea of what is required to develop the uniformisation theory may be had by consulting the presentation in Bieberbach, *Lehrbuch der Funktionentheorie*, 2d ed., vol. 2.

the case of plane regions. We give, namely, in § 10, a proof independent of any previous one of the existence of a conformal map of the region between any two non-intersecting Jordan curves in the plane on a circular ring (Schottky [8]), showing at the same time that the map attaches continuously to a topological correspondence between the boundaries (Osgood and Taylor [9]).

Arbitrary Jordan Contours (with $m(\Gamma_1, \Gamma_2)$ infinite). The sufficient condition (1.5) can be extended so as to apply when any of the quantities $m(\Gamma)$ are not finite. Let

$$(1.8)\quad e(\Gamma_1, \Gamma_2) = m(\Gamma_1) + m(\Gamma_2) - m(\Gamma_1, \Gamma_2) \geqq 0 \quad \text{(by 1.4)};$$

then $e(\Gamma_1, \Gamma_2)$ is defined for all pairs of contours with finite $m(\Gamma_1, \Gamma_2)$. Irrespective of whether $m(\Gamma_1, \Gamma_2)$ is finite or infinite, we then define:

$$(1.9)\qquad \bar{e}(\Gamma_1, \Gamma_2) = \limsup e(\Gamma_1', \Gamma_2'),$$

where Γ_1', Γ_2' are contours with finite $m(\Gamma_1', \Gamma_2')$ tending respectively to Γ_1, Γ_2 in the sense of Fréchet; this means that $d(\Gamma_1, \Gamma_1') \to 0$, $d(\Gamma_2, \Gamma_2') \to 0$, where, considering all possible topological correspondences between Γ and Γ', $d(\Gamma, \Gamma')$ is the lower bound of the greatest distance between corresponding points. Then if

$$(1.10)\qquad \bar{e}(\Gamma_1, \Gamma_2) > 0,$$

the Jordan curves Γ_1, Γ_2 bound a doubly-connected minimal surface; the proof is in § 11.

Interlacing Contours. An interesting and entirely new result that we are able to establish relates to contours that are interlaced. We prove (§§ 9, 11) that *any two Jordan curves which*

[8] F. Schottky, "Ueber die conforme Abbildung mehrfach zusammenhängender ebener Flachen," *Crelle*, vol. 83 (1877), pp. 300–351. Also, P Koebe, "Abhandlungen zur Theorie der konformen Abbildung," *Crelle*, vol. 145 (1915), pp. 177–223.

[9] W. F. Osgood and E. H. Taylor, "Conformal Transformations on the Boundaries of Their Regions of Definition," *Trans. Amer. Math. Soc.*, vol. 14 (1913,, pp. 277–298.

interlace bound a doubly connected minimal surface; of course, this surface is self-intersecting. The result is obtained by showing that the sufficient condition (1.10) is always satisfied by interlacing contours.

Evident Cases of Fulfillment of the Sufficiency Condition. There are certain simple cases where it is obvious from the geometry of the figure that the sufficiency condition (1.5) is satisfied—provided we permit ourselves to use the Lebesgue interpretation of $m(\Gamma)$ in anticipation of proof of its identity with the present one. Consider, for instance, a cylinder or frustum of a cone so flat that the lateral area is less than the sum of the areas of the bases, supposed bounded by rectifiable curves Γ_1, Γ_2. Then, with the Lebesgue interpretation of $m(\Gamma)$, the sufficiency condition is satisfied, and Γ_1, Γ_2 are the boundaries of a doubly-connected minimal surface. More generally, let Γ_1', Γ_2' be any Jordan curves in the planes of Γ_1, Γ_2 respectively, and enclosing these curves. Then the area of the surface formed by joining the ring-shaped regions between Γ_1, Γ_1' and Γ_2, Γ_2' with the cylindrical or conical surface is evidently less than the sum of the plane areas bounded by Γ_1' and Γ_2', so that Γ_1', Γ_2' admit a doubly-connected minimal surface spanned between them.

2. General Formulas

The disposition of the one-contour case in the paper (I) was based on the functional

$$A(g) = \frac{1}{4\pi}\int_0^{2\pi}\int_0^{2\pi} \frac{\sum_{i=1}^{n}[g_i(\theta) - g_i(\varphi)]^2}{4\sin^2\dfrac{\theta-\varphi}{2}}\,d\theta d\varphi,$$

derivable by applying Green's formula to the less explicit definition (1.2). After proving that the minimum of $A(g)$ was attained, the condition $\sum_{i=1}^{n} F_i'^2(w) = 0$, characteristic of a minimal surface, resulted by means of a denumerable infinity of explicitly given variations, expressing the vanishing of the de-

numerable infinity of coefficients in the Taylor expansion of $\sum_{i=1}^{n} F_i^{n}(w)$ about $w = 0$. The analogous method in the two-contour case would be to use the Laurent expansion of $\sum_{i=1}^{n} F_i^{n}(w)$ in the circular ring between C_1 and C_2, and in similar fashion to show that the vanishing of the first variation of $A(g_1, g_2; q)$ implied the vanishing of all the Laurent coefficients. However, this would lead to tedious calculations, and a method is therefore desirable which will give the condition $\sum_{i=1}^{n} F_i^{n}(w) = 0$ *all at once*. Such a method is developed here; we illustrate it first for the one-contour case, recasting our treatment of that into a more elegant form and one more suitable for a generalization of the type we need to the two-contour case.

To solve the Dirichlet problem for any plane region R bounded by a regular contour C, consisting of one or any finite number of pieces, with prescribed boundary values h for the required harmonic function H, we have the formula:

$$H(u, v) = -\frac{1}{2\pi}\int_C h\frac{\partial G}{\partial n}\,ds, \tag{2.1}$$

$G(x, y; u, v)$ denoting the Green's function of the region, vanishing on C and expressible in its interior in the form

$$G(x, y; u, v) = g(x, y; u, v) + \log\sqrt{(x-u)^2 + (y-v)^2},$$

where g is harmonic in x, y and regular throughout the region; the differentiation $\frac{\partial}{\partial n}$ is in the sense of the inner normal.

It is desirable to express $H(u, v)$ as the real part of a function of a complex variable. To that end, we first so express G, considered as a harmonic function of x, y; introducing the notation

$$z = x + iy, \qquad w = u + iv, \qquad w_0 = u - iv,$$

we write

$$G(x, y; u, v) = \Re \log S(z; w, w_0), \tag{2.2}$$

(the notation S and the later notation Z, P are meant to suggest the elliptic functions σ, ζ, $\wp$, to which they are tantamount in the case where the region is a circular ring; the periods of these elliptic functions are 2π and $2i \log q$, q the ratio of the radii).

Suppose C to consist of *any finite number of circles* C_α:

$$C = C_1, \quad \text{or} \quad C = C_1 + C_2, \quad \text{etc.}$$

Then using on each circle polar coordinates referred to its center, $z - a_\alpha = \rho e^{i\theta}$, we have

$$(2.3) \qquad (z - a_\alpha)\frac{\partial}{\partial z}\log S(z; w, w_0) = \rho\frac{\partial G}{\partial \rho} - i\frac{\partial G}{\partial \theta};$$

but on each circular boundary $G = 0$ identically, so that $\dfrac{\partial G}{\partial \theta} = 0$; consequently, with the notation

$$(2.4) \qquad Z(z; w, w_0) = \frac{\partial}{\partial z}\log S(z; w, w_0),$$

we have

$$(2.5) \qquad (z - a_\alpha)Z(z; w, w_0) = \rho\frac{\partial G}{\partial \rho}$$

for z on C_α.

On each circle C_α, $\dfrac{\partial G}{\partial n} = \pm\dfrac{\partial G}{\partial \rho}$ according as C_α is an inner or an outer boundary, so that with $ds = \rho d\theta$ and (2.5), formula (2.1) becomes

$$(2.6) \quad H(u, v) = \frac{1}{2\pi}\int_{C_1} h\cdot(z - a_1)Z(z; w, w_0)d\theta - \frac{1}{2\pi}\int_{C_2} h\cdot(z - a_2)Z(z; w, w_0)d\theta - \text{etc.},$$

the sense of integration on each circle being counter-clock (so as to give positive $d\theta$ and therefore positive ds).

The Green's function $G(x, y; u, v)$ being harmonic in u, v (symmetry property), it is seen from (2.5) that, for z on C_α, the function $(z - a_\alpha)Z(z; u + iv, u - iv)$ is real-valued and harmonic

for u, v in R. This is to say, in terms of w and w_0, that Z obeys the equation

$$\frac{\partial^2 Z}{\partial w \partial w_0} = 0; \tag{2.7}$$

hence [10]

$$Z(z; w, w_0) = Z(z, w) - Z_1(z, w_0), \tag{2.8}$$

the general solution of (2.7), or, with the (for the moment, constant) factor $z - a_\alpha$:

$$(z - a_\alpha)Z(z; w, w_0) = (z - a_\alpha)Z(z, w) - (z - a_\alpha)Z_1(z, w_0). \tag{2.8'}$$

From (2.8) it follows, returning to (2.4) and (2.2), that

$$G(x, y; u, v) = \Re \log \frac{S(z, w)}{S_1(z, w)} \tag{2.9}$$

and

$$Z(z, w) = \frac{\partial}{\partial z} \log S(z, w) \tag{2.10}$$

(the similar formula with the subscript 1 will not be needed).

Now, the Cauchy-Riemann equations for two conjugate harmonic functions $M(u, v)$, $N(u, v)$ are, in terms of the minimal coordinates w, w_0:

$$\frac{\partial M}{\partial w} = \frac{\partial (iN)}{\partial w_0}, \qquad \frac{\partial M}{\partial w_0} = -\frac{\partial (iN)}{\partial w},$$

whereby the conjugate to the real-valued harmonic function (2.8') is, with the factor i annexed:

$$(z - a_\alpha)Z(z, w) + (z - a_\alpha)Z_1(z, w_0) + \gamma(z),$$

the last term entering as constant of integration.

The real-valued harmonic function (2.8') may therefore be expressed as the real part of a function of the complex variable $w = u + iv$ as follows:

$$(z - a_\alpha)Z(z; w, w_0) = \Re\{2(z - a_\alpha)Z(z, w) + \gamma_\alpha(z)\}. \tag{2.11}$$

[10] To avoid multiplying notation we use the same symbol Z to represent the function of z; w, w_0 and also the function of z, w derived from it. The same remark applies to the symbol S in (2.2) and (2.9).

Using this in (2.6), with the observation $d\theta = \dfrac{dz}{i(z - a_\alpha)}$ for z on C_α, we see that the harmonic function $H(u, v)$ may be expressed in the form:

$$H(u, v) = \Re F(w), \tag{2.12}$$

$$F(w) = \frac{1}{\pi i}\int_C h \cdot Z(z, w)dz + c,$$

the sense of integration on each component circle C_α of C being such that the region R is on the left. The constant c is equal to

$$c = \frac{1}{2\pi}\int_C h \cdot \gamma(z)d\theta$$

and may be readily determined in each special case, for instance, by means of the value of H known at a particular point. The function $\gamma(z)$ has different forms $\gamma_\alpha(z)$ on the different circles C_α; it may be taken to be a real-valued function of the complex point z whose locus is C, since a pure imaginary constant added to $F(w)$ evidently does not change $H(u, v)$.

If the boundary value h is a vector function of n components: $g_{1i}(z)$ on C_1, $g_{2i}(z)$ on C_2, etc. $(i = 1, 2, \cdots, n)$, then (2.12) gives for the harmonic surface spoken of in § 1:

$$x_i = H_i(u, v) = \Re F_i(w), \tag{2.13}$$

$$F_i(w) = \frac{1}{\pi i}\sum_\alpha \int_{C_\alpha} g_{\alpha i}(z)Z(z, w)dz + c.$$

From this,

$$F_i'(w) = \frac{1}{\pi i}\sum_\alpha \int_{C_\alpha} g_{\alpha i}(z)P(z, w)dz, \tag{2.14}$$

where

$$P(z, w) = \frac{\partial}{\partial w}Z(z, w) = \frac{\partial^2}{\partial z \partial w}\log S(z, w). \tag{2.15}$$

Hence

$$\sum_{i=1}^{n} F_i'^2(w) = -\frac{1}{\pi^2}\sum_{\alpha\beta}\int_{C_\alpha}\int_{C_\beta}\sum_{i=1}^{n} g_{\alpha i}(z)g_{\beta i}(\zeta) \cdot P(z, w)P(\zeta, w)dzd\zeta, \tag{2.16$_0$}$$

or, in a more geometric form (square of distance) easily seen to be equivalent:

$$(2.16)\quad \sum_{i=1}^{n} F_i'^2(w) = \frac{1}{2\pi^2}\sum_{\alpha\beta}\int_{C_\alpha}\int_{C_\beta}\sum_{i=1}^{n}[g_{\alpha i}(z) - g_{\beta i}(\zeta)]^2 \times P(z, w)P(\zeta, w)dzd\zeta.$$

What has to be observed for this equivalence is that

$$\int_{C_{(w)}} P(z, w)dz = 0,$$

obvious, because the indefinite integral is, by (2.14), the uniform function $S_w(z, w) \div S(z, w)$.

We need finally an expression for the functional $A(g)$ defined in (1.2) explicitly in terms of the boundary (vector) functions $g_\alpha(z)$. To this end, we first shrink the region R slightly to a region $\overline{R}$ by contracting somewhat the outer bounding circle and expanding the inner ones, keeping them respectively concentric to themselves; let the bounding circles of $\overline{R}$ be denoted by $\overline{C}_\beta$. Then, in view of (1.2), we define

$$(2.17)\quad \overline{A}(g) = \frac{1}{2}\int\int_{\overline{R}}\sum_{i=1}^{n}\left[\left(\frac{\partial H_i}{\partial u}\right)^2 + \left(\frac{\partial H_i}{\partial v}\right)^2\right]dudv,$$

and $A(g)$ is the limit of $\overline{A}(g)$ when $\overline{C}_\beta$ tend to C_β.

Green's formula may be applied to (2.17) to express the surface integral as a contour integral:

$$(2.17')\quad \overline{A}(g) = \frac{1}{2}\sum_{\beta}\int_{\overline{C}_\beta}\sum_{i=1}^{n} H_i\frac{\partial H_i}{\partial \rho}\rho d\theta,$$

account being taken of the harmonic character of $H_i(u, v)$. By a formula analogous to (2.3), viz.

$$\rho\frac{\partial H_i}{\partial \rho} = (w - a_\beta)F_i'(w) + i\frac{\partial H_i}{\partial \theta},$$

and

$$d\theta = \frac{dw}{i(w - a_\beta)},$$

this becomes

$$\overline{A}(g) = \frac{1}{2i}\sum_{\beta}\int_{\overline{C}_\beta}\sum_{i=1}^{n} H_i(u, v)F_i'(w)dw + \frac{i}{2}\sum_{\beta}\int_{\overline{C}_\beta}\sum_{i=1}^{n} H_i\frac{\partial H_i}{\partial\theta}d\theta.$$

The indefinite integral for the second term being the uniform function $\frac{1}{2}\sum_{i=1}^{n} H_i^2$, this term vanishes; and then using the expression (2.14) for $F_i'(w)$ together with a device like that employed to pass from (2.16_0) to (2.16), we have

$$\overline{A}(g) = \frac{1}{4\pi}\sum_{\alpha\beta}\int_{C_\alpha}\int_{\overline{C}_\beta}\sum_{i=1}^{n}[g_{\alpha i}(z) - H_i(u, v)]^2P(z, w)dzdw.$$

If now we perform formally the limit process $\overline{C}_\beta \to C_\beta$, then $w = u + iv$ describing $\overline{C}_\beta$ approaches to ζ describing C_β, and $H_i(u, v)$ to $g_{\beta i}(\zeta)$, so that we get

$$(2.18)\quad A(g) = \frac{1}{4\pi}\sum_{\alpha\beta}\int_{C_\alpha}\int_{C_\beta}\sum_{i=1}^{n}[g_{\alpha i}(z) - g_{\beta i}(\zeta)]^2P(z, \zeta)dzd\zeta;$$

and there is, in fact, no essential difficulty in establishing the function-theoretic validity of this limit process.[11] It must be observed that the double integral expressing $A(g)$ is improper, since the factor $P(z, \zeta)$, as we shall see in formulas (3.4) and (4.6), becomes infinite like $\frac{1}{(z-\zeta)^2}$ for $z = \zeta$. The integral is defined, of course, by integrating first with $|z - \zeta| > \epsilon$ and then taking the limit as $\epsilon \to 0$. The integral for $\overline{A}(g)$, on the other hand, is proper, since the circles $\overline{C}_\beta$ have no point in common with the circles C_α, and therefore in $\overline{A}(g)$ always $z \neq w$.

3. The Case of a Single Contour

We shall now apply the formulas just developed to the case of a single contour.

Taking C to be the unit circumference, the Green's function

[11] Cf. (I), § 21.

for R, the unit circular disc, is seen to be

$$(3.1) \qquad G(x, y; u, v) = \Re \log \frac{z - w}{zw_0 - 1} = \log \frac{|z - w|}{|zw_0 - 1|}$$

(notation: the subscript 0 denotes the conjugate complex quantity), this evidently having all the requisite properties: logarithmically singular at $z = w$, regular harmonic everywhere else, vanishing for $|z| = 1 \left(z = \frac{1}{z_0} \right)$.

Hence, by (2.9), (2.10) and (2.15):

$$(3.2) \qquad S(z, w) = \log (z - w),$$

$$(3.3) \qquad Z(z, w) = \frac{1}{z - w},$$

$$(3.4) \qquad P(z, w) = \frac{1}{(z - w)^2}.$$

By (2.13), the harmonic surface with the boundary parametrically represented by $x_i = g_i(z)$ is

$$(3.5) \qquad x_i = \Re F_i(w),$$

$$F_i(w) = \frac{1}{\pi i} \int_C g_i(z) \frac{dz}{z - w} + c.$$

The constant c can be determined by putting $w = 0$; according to the mean value property of harmonic functions we have

$$\Re F_i(0) = \frac{1}{2\pi} \int_C g_i(\theta) d\theta = \frac{1}{2\pi i} \int_C g_i(z) \frac{dz}{z},$$

and we may as well write

$$F_i(0) = \frac{1}{2\pi i} \int_C g_i(z) \frac{dz}{z},$$

since $F_i(w)$ is determined only up to an additive pure imaginary constant. This gives, then,

$$c = -\frac{1}{2\pi i} \int_C g_i(z) \frac{dz}{z},$$

and therefore

$$(3.6)\qquad F_i(w) = \frac{1}{\pi i}\int_C g_i(z)\frac{dz}{z-w} - \frac{1}{2\pi i}\int_C g_i(z)\frac{dz}{z} = \frac{1}{2\pi i}\int_C g_i(z)\frac{z+w}{z-w}\frac{dz}{z}.$$

The formulas (2.16) and (2.18) take the form

$$(3.7)\qquad \sum_{i=1}^{n} F_i'^2(w) = \frac{1}{2\pi^2}\int_C\int_C \sum_{i=1}^{n}[g_i(z) - g_i(\zeta)]^2 \frac{dzd\zeta}{(z-w)^2(\zeta-w)^2},$$

$$(3.8)\qquad A(g) = \frac{1}{4\pi}\int_C\int_C \sum_{i=1}^{n}[g_i(z) - g_i(\zeta)]^2 \frac{dzd\zeta}{(z-\zeta)^2}.$$

It will be noticed that, although complex quantities enter into the expression for $A(g)$, the value of that functional is always positive real.

The properties of $A(g)$ can be derived exactly as in our first paper: its invariance under linear fractional transformation of z (generally converting the unit circle into a different circle), its lower semi-continuity—as limit of the continuous functional $\bar{A}(g)$, which tends to it in increasing because of the positive integrand $\frac{1}{2}(E+G)$: hence, by Fréchet's thesis, the minimum of $A(g)$ when all parametric representations of the given contour are considered is attained for a certain representation

$$x_i = g_i^*(z),$$

which can be proved to be a "proper" representation, as in the cited paper.

The point is to show that the vanishing of the first variation of $A(g)$ for $g = g^*$ leads to the condition for a minimal surface:

$$\sum_{i=1}^{n} F_i'^2(w) = 0.$$

This we easily do by means of the special variation

$$(3.9)\qquad z' = z + \frac{\lambda}{z-w},$$

w denoting any fixed point interior to the unit circle. Now this is not, without further discussion, an admissible variation, since it does not convert the unit circle $|z| = 1$ into itself at all, let alone in a monotonic continuous way; but if the reader will follow us through the next calculations, we will readily justify the use of this variation immediately afterwards.

We drop all terms beyond the first power of λ:

$$z' = z + \frac{\lambda}{z - w}, \qquad \zeta' = \zeta + \frac{\lambda}{\zeta - w};$$

$$dz' = \left\{ 1 - \frac{\lambda}{(z - w)^2} \right\} dz, \qquad d\zeta' = \left\{ 1 - \frac{\lambda}{(\zeta - w)^2} \right\} d\zeta;$$

$$dz'd\zeta' = \left\{ 1 - \frac{\lambda}{(z - w)^2} - \frac{\lambda}{(\zeta - w)^2} \right\} dzd\zeta,$$

$$z' - \zeta' = (z - \zeta) \left\{ 1 - \frac{\lambda}{(z - w)(\zeta - w)} \right\},$$

$$\frac{dz'd\zeta'}{(z' - \zeta')^2} = \frac{dzd\zeta}{(z - \zeta)^2} \left[1 - \lambda \left\{ \frac{1}{(z - w)^2} + \frac{1}{(\zeta - w)^2} - \frac{2}{(z - w)(\zeta - w)} \right\} \right].$$

The coefficient of λ is

$$\left(\frac{1}{z - w} - \frac{1}{\zeta - w} \right)^2$$

and we have—what is the fundamental algebraic identity of the calculations:

$$\frac{1}{z - w} - \frac{1}{\zeta - w} = - \frac{z - \zeta}{(z - w)(\zeta - w)}; \tag{3.10}$$

thus

$$\frac{dz'd\zeta'}{(z' - \zeta')^2} = \frac{dzd\zeta}{(z - \zeta)^2} - \lambda \frac{dzd\zeta}{(z - w)^2(\zeta - w)^2}. \tag{3.11}$$

We apply the transformation (3.9) to $A(g^*)$, and expand in powers of λ. According to (3.11), this expansion begins:

$$A(g^*) - \lambda \cdot \frac{1}{4\pi} \int_C \int_C \sum_{i=1}^{n} [g_i^*(z) - g_i^*(\zeta)]^2 \frac{dzd\zeta}{(z - w)^2(\zeta - w)^2}; \tag{3.12}$$

but here *the coefficient of* λ *is, apart from a numerical factor, the same as the expression* (3.7) *for* $\sum_{i=1}^{n} F_i'^2(w)$, so that the vanishing of that coefficient gives

$$\sum_{i=1}^{n} F_i'^2(w) = 0. \tag{3.13}$$

It remains only to justify the use of the variation (3.9). Now, while this is not itself an admissible variation, it is expressible as a linear combination of admissible variations.

If w denote any fixed point interior to the unit circle, z a point moving over the unit circumference, and λ a real parameter, then

$$\frac{1}{z(z-w)} \quad \text{and} \quad \frac{1}{z_0(z_0-w_0)}$$

are conjugate imaginary, so that

$$\lambda\left\{\frac{1}{z(z-w)}-\frac{1}{z_0(z_0-w_0)}\right\} \quad \text{and} \quad i\lambda\left\{\frac{1}{z(z-w)}+\frac{1}{z_0(z_0-w_0)}\right\}$$

are pure imaginary. Hence the transformations

$$z' = ze^{\lambda\left\{\frac{1}{z(z-w)}-\frac{1}{z_0(z_0-w_0)}\right\}} \tag{3.14$_1$}$$

$$z' = ze^{i\lambda\left\{\frac{1}{z(z-w)}+\frac{1}{z_0(z_0-w_0)}\right\}} \tag{3.14$_2$}$$

convert points on the unit circumference again into such points. Furthermore, for $\lambda = 0$ these transformations reduce to the identity, for which $\frac{dz'}{dz} = 1$ identically. Therefore, on account of the uniform continuity of all the functions involved in λ and z (w fixed), $\frac{dz'}{dz}$ stays positive for all z on C if $|\lambda|$ is small enough, and (3.14) therefore represents for $|\lambda|$ sufficiently small a monotonic continuous transformation of the unit circumference into itself.

If the transformation (3.14) is applied to the minimizing representing representation $x_i = g_i^*(z)$, then $A(g)$ becomes a

function of λ, $A_1(\lambda)$, having at $\lambda = 0$ an absolute minimum with respect to a certain interval about that value. $A_1(\lambda)$ is, furthermore, representable as a power series in λ:

$$A_1(\lambda) = A(g^*) + \lambda V_1 + \cdots,$$

obtained by formal expansion after the substitution (3.14_1) for z', ζ'; the function-theoretic justification is by easy application of classic theorems (Cauchy appraisal formula, Weierstrass double-series theorem), see (I), §§ 13, 14. From the minimum property of $A_1(\lambda)$ follows the vanishing of its first derivative:

$$V_1 = 0. \tag{3.15$_1$}$$

Using similarly the transformation (3.14_2), we get

$$A_2(\lambda) = A(g^*) + i\lambda V_2 + \cdots,$$

and consequently

$$V_2 = 0. \tag{3.15$_2$}$$

Now the transformations (3.14) may be expanded in power series of λ:

$$\begin{aligned} z' &= z + \lambda\left\{\frac{1}{z-w} - \frac{z}{z_0(z_0-w_0)}\right\} + \cdots, \\ z' &= z + i\lambda\left\{\frac{1}{z-w} + \frac{z}{z_0(z_0-w_0)}\right\} + \cdots, \end{aligned} \tag{3.16}$$

and evidently, as far as the formal calculations are concerned, the same values would be obtained for V_1, V_2 if we broke off (3.16) after the first power of λ:

$$\begin{aligned} z' &= z + \lambda\left\{\frac{1}{z-w} - \frac{z}{z_0(z_0-w_0)}\right\}, \\ z' &= z + i\lambda\left\{\frac{1}{z-w} + \frac{z}{z_0(z_0-w_0)}\right\} \end{aligned}$$

(infinitesimal transformations).

As is well-known, any linear combination of these infinitesimal transformations will give for the formal variation the same linear

combination of V_1 and V_2; we choose as linear combination one-half the sum:

$$z' = z + \frac{\lambda}{z - w},$$

which will give as formal variation

$$\tfrac{1}{2}(V_1 + V_2),$$

and by (3.15_1) and (3.15_2) we have the right to put this equal to zero.

4. Formulas for the Two-contour Case

The region R on which the minimal surface will be represented conformally is here a circular ring, with bounding circles C_1 and C_2 of radii 1 and q, $0 < q < 1$.

The first step is to find the Green's function for this region.[12] We write, as in (2.2):

$$(4.1) \quad G(x, y; u, v) = \Re \log S(z; w, w_0) = \log |S(z; w, w_0)|,$$

and $|S(z)|$ (to emphasize only its dependence on z) must be identically equal to unity on C_1, C_2. It follows by the Schwarz symmetry principle, since $S(z)$ thus converts the circles C_1, C_2 in the z-plane into the unit circle in the S-plane, that points inverse to the circle $C_1\left(z, \frac{1}{z_0}\right)$ or to the circle $C_2\left(z, \frac{q^2}{z_0}\right)$ give values of S inverse to the unit circle $\left(S, \frac{1}{S_0}\right)$, i.e. we have the functional equations

$$(4.2) \qquad S(z)\overline{S\left(\frac{1}{z_0}\right)} = 1, \qquad S(z)\overline{S\left(\frac{q^2}{z_0}\right)} = 1,$$

the bar as well as the subscript 0 denoting the conjugate complex quantity.

[12] The developments here follow Courant-Hilbert, *Methoden der Mathematischen Physik*, I, chap. V, § 12, no. 10. The statement that the formula there derived is valid also for complex values of c, the "Aufpunkt," is a slip; the conjugate complex quantity c_0 must intervene.

Combining the inversions in C_1 and C_2, we see that S must therefore take the same value at all the points

$$z,\ q^{\pm 2}z,\ q^{\pm 4}z,\ \cdots$$

and the value $\dfrac{1}{S_0}$ at the points

$$\frac{1}{z_0},\ \frac{q^{\pm 2}}{z_0},\ \frac{q^{\pm 4}}{z_0},\ \cdots.$$

Now to give G the requisite logarithmic singularity at w, the function $S(z)$ must have a zero at that point, and therefore zeros at all the points

$$w,\ q^{\pm 2}w,\ q^{\pm 4}w,\ \cdots,$$

and poles at all the points

$$\frac{1}{w_0},\ \frac{q^{\pm 2}}{w_0},\ \frac{q^{\pm 4}}{w_0},\ \cdots.$$

A function with these zeros and poles is

$$F(z) = \frac{\left(1 - \dfrac{z}{w}\right)\prod_{m=1}^{\infty}\left(1 - \dfrac{q^{2m}z}{w}\right)\left(1 - \dfrac{q^{2m}w}{z}\right)}{(1 - w_0 z)\prod_{m=1}^{\infty}(1 - q^{2m}w_0 z)\left(1 - \dfrac{q^{2m}}{w_0 z}\right)}.$$

Trying to see how far $F(z)$ obeys the functional equations (4.2), we find

$$F(z)\overline{F\left(\frac{1}{z_0}\right)} = \frac{1}{ww_0}, \qquad F(z)\overline{F\left(\frac{q^2}{z_0}\right)} = 1.$$

We can easily modify $F(z)$ with a factor az^b, a and b real:

$$S(z) = az^b F(z),$$

so that the functional equations (4.2) will be verified. The calculations are:

$$S(z)\overline{S\left(\frac{1}{z_0}\right)} = az^b \cdot \frac{a}{z^b} \cdot F(z)\overline{F\left(\frac{1}{z_0}\right)} = \frac{a^2}{ww_0},$$

$$S(z)\overline{S\left(\frac{q^2}{z_0}\right)} = az^b \cdot \frac{aq^{2b}}{z^b} \cdot F(z)\overline{F\left(\frac{q^2}{z_0}\right)} = a^2q^{2b};$$

so that we have for the determination of a, b the equations

$$\frac{a^2}{ww_0} = 1, \qquad a^2q^{2b} = 1,$$

giving

$$a = (ww_0)^{1/2}, \qquad b = -\frac{\log ww_0}{2\log q}.$$

Hence the Green's function (4.1) for the circular ring is

(4.3) $G(x, y; u, v)$

$$= \Re \log \frac{w^{1/2}z^{-\frac{\log w}{2\log q}}\left(1-\frac{z}{w}\right)\prod_{m=1}^{\infty}\left(1-\frac{q^{2m}z}{w}\right)\left(1-\frac{q^{2m}w}{z}\right)}{w_0^{-1/2}z^{\frac{\log w_0}{2\log q}}(1-w_0z)\prod_{m=1}^{\infty}(1-q^{2m}w_0z)\left(1-\frac{q^{2m}}{w_0z}\right)},$$

of the predicted form (2.9):

$$G(x, y; u, v) = \Re \log \frac{S(z, w)}{S_1(z, w_0)},$$

with

(4.4) $$S(z, w) = w^{1/2}z^{-\frac{\log w}{2\log q}}\left(1-\frac{z}{w}\right)\prod_{m=1}^{\infty}\left(1-\frac{q^{2m}z}{w}\right)\left(1-\frac{q^{2m}w}{z}\right)$$

(we shall not need $S_1(z, w_0)$).

By the definitions (2.10), (2.15):

(4.5) $$Z(z, w) = -\frac{\log w}{2\log q}\cdot\frac{1}{z} + \sum_{m=0}^{\infty}\frac{q^{2m}}{q^{2m}z - w} + \sum_{m=1}^{\infty}\left(\frac{1}{z-q^{2m}w} - \frac{1}{z}\right),$$

(4.6) $$P(z, w) = -\frac{1}{2\log q}\cdot\frac{1}{zw} + \sum_{m=-\infty}^{\infty}{}'\frac{q^{2m}}{(q^{2m}z - w)^2} = P(w, z).$$

The harmonic surface with the boundaries

$$x_i = g_{1i}(z), \qquad x_i = g_{2i}(z)$$

is, by (2.13),

(4.7) $$x_i = \Re F_i(w),$$

$$F_i(w) = \frac{1}{\pi i}\int g_{1i}(z)Z(z, w)dz + \frac{1}{\pi i}\int_C g_{2i}(z)Z(z, w)dz + c_i,$$

and the constant c_i can be determined by using the Laurent expansion and equating constant terms on both sides (it will be recalled that an additive pure imaginary constant may be neglected in $F_i(w)$, since that has no effect on $\Re F_i(w)$); we thus find

$$c_i = -\frac{1}{2\pi i}\int_{C_1} g_{1i}(z)\,\frac{dz}{z}.$$

We repeat also the formulas (2.18), (2.16), where to abbreviate we leave implicit the dependence of $P(z, \zeta)$ on q:

$$P(z, \zeta) = P(z, \zeta; q).$$

$$(4.8)\quad A(g_1, g_2; q) = \frac{1}{4\pi}\left\{\int_{C_1}\int_{C_1}\sum_{i=1}^{n}[g_{1i}(z) - g_{1i}(\zeta)]^2 P(z, \zeta)dzd\zeta + 2\int_{C_1}\int_{C_2}\sum_{i=1}^{n}[g_{1i}(z) - g_{2i}(\zeta)]^2 P(z, \zeta)dzd\zeta + \int_{C_2}\int_{C_2}\sum_{i=1}^{n}[g_{2i}(z) - g_{2i}(\zeta)]^2 P(z, \zeta)dzd\zeta\right\}.$$

$$(4.9)\quad \sum_{i=1}^{n} F_i'^2(w) = \frac{1}{2\pi^2}\left\{\int_{C_1}\int_{C_1}\sum_{i=1}^{n}[g_{1i}(z) - g_{1i}(\zeta)]^2 P(z, w)P(\zeta, w)dzd\zeta + 2\int_{C_1}\int_{C_2}\sum_{i=1}^{n}[g_{1i}(z) - g_{2i}(\zeta)]^2 P(z, w)P(\zeta, w)dzd\zeta + \int_{C_2}\int_{C_2}\sum_{i=1}^{n}[g_{2i}(z) - g_{2i}(\zeta)]^2 P(z, w)P(\zeta, w)dzd\zeta\right\}.$$

The sense of integration is such that the interior of the circular ring is on the left: counter-clockwise on C_1, clockwise on C_2.

5. The Functional A in Terms of Fourier Coefficients

We shall find useful the expression of $A(g_1, g_2; q)$ in terms of the Fourier coefficients of g_1 and g_2.

For the harmonic vector function $x_i = \Re F_i(w)$ $(i = 1, 2, \cdots, n)$ we have the Laurent expansion in the circular ring:

$$(5.1)\quad \Re F(w) = H(\rho, \theta) = \frac{\alpha}{-\log q}\log \rho + A_0 + \sum_{m=1}^{\infty} \rho^m (A_m \cos m\theta + B_m \sin m\theta) + \sum_{m=1}^{\infty} \rho^{-m}(A_{-m}\cos m\theta - B_{-m}\sin m\theta),$$

where, if $a_{1,m}$, $b_{1,m}$ (vectors with n components: a_{1im}, b_{1im}, $i = 1, 2, \cdots, n$) denote the Fourier coefficients of $g_1(\theta)$ and $a_{2,m}$, $b_{2,m}$ of $g_2(\theta)$, we have

$$(5.2)\qquad A_m = \frac{a_{1,m} - q^m a_{2,m}}{1 - q^{2m}}, \qquad A_{-m} = \frac{a_{1,m} - q^{-m}a_{2,m}}{1 - q^{-2m}};$$

similarly for B_m,

$$\alpha = \frac{1}{2\pi}\int_0^{2\pi} g_1(\theta)d\theta - \frac{1}{2\pi}\int_0^{2\pi} g_2(\theta)d\theta,$$

$$A_0 = \frac{1}{2\pi}\int_0^{2\pi} g_1(\theta)d\theta.$$

We have by (2.17′), both circles of integration being sensed counter-clockwise,

$$(5.3)\quad A(g_1, g_2; q) = \lim_{\rho\to 1}\frac{1}{2}\int_{C_\rho} H\frac{\partial H}{\partial \rho}\rho d\theta - \lim_{\rho\to q}\frac{1}{2}\int_{C_\rho} H\frac{\partial H}{\partial \rho}\rho d\theta = \frac{1}{4}\lim_{\rho\to 1}\rho\frac{\partial}{\partial\rho}\int_{C_\rho} H^2 d\theta - \frac{1}{4}\lim_{\rho\to q}\rho\frac{\partial}{\partial\rho}\int_{C_\rho} H^2 d\theta.$$

If in (5.1) ρ be regarded as fixed, then that formula represents a Fourier expansion, and by Parseval's theorem:

$$\int_0^{2\pi} H^2 d\theta = 2\pi\left(\frac{\alpha}{-\log q}\log\rho + A_0\right)^2 + \pi\sum_{m=1}^{\infty}[(A_m^2 + B_m^2)\rho^{2m} + (A_{-m}^2 + B_{-m}^2)\rho^{-2m} + 2(A_mA_{-m} - B_mB_{-m})].$$

Substituting in (5.3):

$$(5.4)\quad A(g_1, g_2; q) = \frac{4\pi\alpha^2}{-\log q} + \frac{\pi}{2}\sum_{m=1}^{\infty} m \frac{\begin{array}{r}(1+q^{2m})(a_{1,m}^2 + b_{1,m}^2 + a_{2,m}^2 + b_{2,m}^2) \\ -4q^m(a_{1,m}a_{2,m} + b_{1,m}b_{2,m})\end{array}}{1-q^{2m}},$$

a quadratic form in the infinitude of Fourier coefficients, which when $q = 0$ reduces to that found in (I), § 21 for $A(g)$. It is understood that the squares and products above are scalar products of vectors:

$$(5.5)\qquad a_{1,m}^2 = \sum_{i=1}^{n} a_{1im}^2, \qquad a_{1,m}a_{2,m} = \sum_{i=1}^{n} a_{1im}a_{2im}.$$

The formula (5.4) is valuable in enabling us to study the functional A in the neighborhood of $q = 0$ and $q = 1$.

In the neighborhood of $q = 0$, we get an expansion consisting of the reciprocal of a logarithm followed by a power series:

$$(5.6)\quad A(g_1, g_2; q) = A(g_1) + A(g_2) + \frac{4\pi\alpha^2}{-\log q} - 2\pi q(a_{1,1}a_{2,1} + b_{1,1}b_{2,1}) + \cdots;$$

evidently in case $\alpha \neq 0$ the (positive) term $\dfrac{4\pi\alpha^2}{-\log q}$ is dominant for q near to zero.

Writing $q = 1 - \epsilon$, we find that, asymptotically in the neighborhood of $q = 1$,

$$(5.7)\quad A(g_1, g_2; q) \sim \frac{3\pi\alpha^2 + \frac{1}{2}\int_0^{2\pi}\sum_{i=1}^{n}[g_{1i}(\theta) - g_{2i}(\theta)]^2 d\theta}{1-q};$$

moreover, this asymptotic expression holds *uniformly* for $q \to 1$ provided the contours Γ_1, Γ_2 remain in some fixed finite region of space. If $d > 0$ denote the (minimum) distance between Γ_1 and Γ_2, we have for the numerator the appraisal

$$(5.7')\qquad \geqq \pi d^2.$$

We shall need too the asymptotic expression derived by differentiation of (5.7) as to q: for $q \to 1$,

$$(5.8)\qquad \frac{\partial A(g_1, g_2; q)}{\partial q} \sim \frac{3\pi\alpha^2 + \frac{1}{2}\int_0^{2\pi} \sum_{i=1}^{n} [g_{1i}(\theta) - g_{2i}(\theta)]^2 d\theta}{(1-q)^2},$$

and the same remark as to uniformity applies here.

6. The Minimizing Representation

Suppose given the two contours Γ_1, Γ_2, Jordan curves not intersecting one another, with finite $m(\Gamma_1, \Gamma_2)$, and

$$(6.1)\qquad m(\Gamma_1, \Gamma_2) < m(\Gamma_1) + m(\Gamma_2).$$

We shall term the system

$$(6.2)\qquad [g_{1i}, g_{2i}; q]$$

a *representation* of the contours,[13] putting thus in evidence their parametric equations

$$x_i = g_{1i}(z), \qquad x_i = g_{2i}(z),$$

as well as q, the ratio of the radii of the concentric circles C_1, C_2 described by the complex variable z.

It was shown in (I) that the totality of parametric representations of a given contour is compact provided we introduce improper and degenerate representations as follows:

1° improper of the first kind, a point of C corresponds to an arc of Γ less than all of Γ;

2° improper of the second kind, an arc of C less than all of C corresponds to a point of Γ;

3° degenerate, all of C corresponds to a point of Γ and all of Γ to a point of C.

Graphically, with a C and Γ axis, the proper representations are depicted by properly monotonic continuous curves; type 1° has discontinuities with unequal left and right limits and a

[13] Either sense of description may be fixed on Γ_1, Γ_2 and the theory carried through separately for each choice (there are essentially two).

finite jump between these limits, representing the arc of Γ to which the point of discontinuity on C corresponds; type 2° has intervals of constancy, representing the arcs of C which correspond to single points of Γ; type 3° is represented by a parallel to the C axis together with a parallel to the Γ axis.

The range of q, $0 < q < 1$, is not compact, but becomes so if we include the extreme values: $0 \leqq q \leqq 1$.

As extended by the above conventions, the composite elements (6.2) form a compact set.

We must next extend to the improper representations the definitions of $A(g_1, g_2; q)$. First we see that if $q = 0$, $P(z, \zeta)$ reduces to $\dfrac{1}{(z-\zeta)^2}$, and (4.8) then becomes, with reference to (3.8):

$$A(g_1, g_2; 0) = A(g_1) + A(g_2), \tag{6.3}$$

the integral over C_1C_2 vanishing since C_2 reduces to a point (whereas the integral over C_2C_2 is independent of the radius of C_2).

For $q \to 1$ we have the asymptotic expression (5.7) with the appraisal (5.7′), and therefore

$$A(g_1, g_2; 1) = +\infty. \tag{6.4}$$

For an improper representation of the first kind we have $A(g_1, g_2; q) = +\infty$, for then the integrand in the integral over C_1C_1 (or C_2C_2) is, in the neighborhood of the point of discontinuity, asymptotically equal to

$$\frac{l^2}{(z-\zeta)^2},$$

where $l > 0$ is the length of the chord of discontinuity, and the order to which this becomes infinite when $z = \zeta$ implies the value $+\infty$ for the integral.

In a degenerate representation the arc of discontinuity is the whole contour, e.g. Γ_2, and its chord l is zero, so that the preceding argument no longer applies. There is, however, little trouble in seeing that if we define $A(g_1, g_2; q)$ for degenerate representations by lower semi-continuity, i.e. as limit inferior of the values

of A for approaching proper representations; then for a representation degenerate as to Γ_2:

$$A(g_1, g_2; q) > A(g_1) + min\, A(g_2), \tag{6.5}$$

whence

$$A(g_1, g_2; q) > m(\Gamma_1) + m(\Gamma_2). \tag{6.5'}$$

Proof. If $g_{2i}(z)$ tends towards a point a_i of Γ_2, then in (4.8), in the integral over C_2C_2, only the term $\dfrac{1}{(z-\zeta)^2}$ in $P(z, \zeta)$ is effective, and gives as limit inferior $min\, A(g_2)$, since in the neighborhood of a degenerate representation the functional $A(g_2)$ takes its whole stock of values—for by linear fractional transformation of C_2, not affecting $A(g_2)$, any representation can be brought arbitrarily close to a degenerate one ((I), § 6).

The remaining terms in (4.8) evidently represent, according to the derivation of that formula, one-half the Dirichlet integral formed for the harmonic vector function in the circular ring whose boundary values are

$$g_{1i}(z) \text{ on } C_1, \qquad a_i \text{ on } C_2.$$

If this function be prolonged as the constant a_i into the interior of C_2, then we have a piece-wise analytic function in the interior of C_1, and the Dirichlet integral for this function is the same as for the original function in the ring C_1C_2, since the contribution of the constant part is zero. Now the Dirichlet integral for the function of two pieces is, by a well-known theorem, greater than for the function harmonic throughout the interior of C_1 with the boundary values $g_{1i}(z)$, i.e. greater than $A(g_1)$. Consequently we have (6.5), and therefrom (6.5').

A similar argument shows that if the representation is doubly degenerate, into a point a_i for Γ_1 and b_i for Γ_2 with distance l necessarily > 0, then

$$A(g_1, g_2; q) = \min A(g_1) + \min A(g_2) + \frac{\pi}{-\log q} l^2 > m(\Gamma_1) + m(\Gamma_2). \tag{6.6}$$

Summarizing, we see that for $q = 0$, $q = 1$, and for all improper representations of types 1° and 3°, we have

$$A(g_1, g_2; q) \geqq m(\Gamma_1) + m(\Gamma_2) > m(\Gamma_1, \Gamma_2). \tag{6.7}$$

It is impossible therefore that the minimum value $m(\Gamma_1, \Gamma_2)$ *of* $A(g_1, g_2; q)$ *be attained for* $q = 0$, *or* $q = 1$, *or for any improper representation of type* 1° *or* 3°.

It remains to exclude the improper representations of type 2°. This is easily done if we permit ourselves to anticipate the proof in the next section (no vicious circle) that the vanishing of the first variation of $A(g_1, g_2; q)$ leads to the condition

$$\sum_{i=1}^{n} F_i'^2(w) = 0.$$

For under that condition it can be shown with Schwarzian symmetry as in (I), § 18, that the vector function $\Re F_i(w)$ cannot reduce to a constant on an arc of C_1 or C_2 without reducing identically to a constant: now this cannot be, for Γ_1 and Γ_2 have no point in common.

That the functional $A(g_1, g_2; q)$ is lower semi-continuous results immediately from its expression (1.2) as an integral with positive integrand. It follows by Fréchet's thesis that the minimum of $A(g_1, g_2; q)$ on the compact closed set $[g_1, g_2; q]$ is attained; and we have seen that this cannot be for an improper representation. Therefore *there exists a proper minimizing representation*

$$x_i = g_{1i}^*(z), \qquad x_i = g_{2i}^*(z)$$

with a value q^* of q, > 0 and < 1.

7. Vanishing of the First Variation of $A(g_1, g_2; q)$

We have now to show that the vanishing of the first variation of A for the minimizing representation gives the condition

$$\sum_{i=1}^{n} F_i'^2(w) = 0$$

characteristic of a minimal surface.

Playing a central role in the calculation is the following identity:

$$(7.1)\qquad \begin{aligned} w^2P(z,w)P(\zeta,w) = &-P(z,\zeta)\{z^2Z_z(z,w)+\zeta^2Z_\zeta(\zeta,w)\} \\ &+z\zeta\{Z(z,w)P_\zeta(z,\zeta)+Z(\zeta,w)P_z(z,\zeta)\} \\ &+\frac{1}{2}q\frac{\partial}{\partial q}P(z,\zeta), \end{aligned}$$

whose proof we postpone to the next section. This will here take the place of the much simpler identity (3.10):

$$\frac{1}{z-w}-\frac{1}{\zeta-w}=-\frac{z-\zeta}{(z-w)(\zeta-w)}$$

of the one-contour case. It amounts to the Weierstrass partial differential equation for the elliptic function $\wp(z;\ \omega_1,\ \omega_2)$ considered as depending on all three arguments.[14]

We apply to the minimizing representation the special variation

$$(7.2)\qquad z' = z + \lambda z^2 Z(z,w),$$

whose use may be justified exactly as was that of (3.9)

$$z' = z + \frac{\lambda}{z-w}$$

in the one-contour case. Making the substitution (7.2) and the same substitution for ζ in the expression (4.8) for $A(g_1,\ g_2;\ q)$, then expanding in powers of λ and equating the coefficient of the first power of λ to zero, we get

$$\begin{aligned} (7.3)\quad 0 = \sum_{\alpha\beta}\int_{C_\alpha}\int_{C_\beta}\sum_{i=1}^{n}&[g_{\alpha i}^*(z)-g_{\beta i}^*(\zeta)]^2 \\ &\times[P(z,\zeta)\{z^2Z_z(z,w)+\zeta^2Z_\zeta(\zeta,w)\} \\ &+2P(z,\zeta)\{zZ(z,w)+\zeta Z(\zeta,w)\} \\ &+z^2P_z(z,\zeta)Z(z,w)+\zeta^2P_\zeta(z,\zeta)Z(\zeta,w)]dzd\zeta. \end{aligned}$$

[14] See C. Jordan, *Cours d'Analyse*, 2d ed., vol. 2, p. 457.

Denoting by [ ] the expression in the square bracket, we have by addition with the identity (7.1):

$$w^2P(z, w)P(\zeta, w) + [\ ]$$
$$= zZ(z, w)\{zP_z(z, \zeta) + 2P(z, \zeta) + \zeta P_\zeta(z, \zeta)\}$$
$$+\zeta Z(\zeta, w)\{zP_z(z, \zeta) + 2P(z, \zeta) + \zeta P_\zeta(z, \zeta)\}$$
$$+\frac{1}{2}q\frac{\partial}{\partial q}P(z, \zeta).$$

By the homogeneity of the minus second degree of $P(z, \zeta)$, the curved brackets vanish:

$$zP_z(z, \zeta) + \zeta P_\zeta(z, \zeta) = -2P(z, \zeta),$$

so that

$$w^2P(z, w)P(\zeta, w) = -[\ ] + \frac{1}{2}q\frac{\partial}{\partial q}P(z, \zeta). \tag{7.4}$$

Now, besides (7.3), expressing the vanishing of the first variation when the parametric representations of Γ_1 and Γ_2 are varied, we have also the vanishing of the partial derivative as to q of $A(g_1, g_2; q)$ for the minimizing representation, i.e.

$$0 = \sum_{\alpha\beta}\int_{C_\alpha}\int_{C_\beta}\sum_{i=1}^{n}[g_{\alpha i}^*(z) - g_{\beta i}^*(\zeta)]^2\frac{\partial}{\partial q}P(z, \zeta)dzd\zeta \tag{7.5}$$

for $q = q^*$.

Combining (7.3) and (7.5) by means of (7.4), we obtain

$$0 = \sum_{\alpha\beta}\int_{C_\alpha}\int_{C_\beta}\sum_{i=1}^{n}[g_{\alpha i}^*(z) - g_{\beta i}^*(\zeta)]^2P(z, w)P(\zeta, w)dzd\zeta, \tag{7.6}$$

and by formula (4.9) this is the same as

$$\sum_{i=1}^{n}F_i'^2(w) = 0.$$

8. Proof of the Fundamental Identity

To complete the argument it remains only to prove the fundamental identity (7.1). For this purpose we start with

$$\frac{w}{(w - q^{2m}z)(w - q^{2n}\zeta)} = \frac{1}{q^{2m}z - q^{2n}\zeta}\left\{\frac{q^{2m}z}{w - q^{2m}z} - \frac{q^{2n}\zeta}{w - q^{2n}\zeta}\right\}.$$

Squaring, we get

$$\frac{w^3}{(w-q^{2m}z)^2(w-q^{2n}\zeta)^2} = \frac{1}{(q^{3m}z-q^{2n}\zeta)^2}\left\{\frac{q^{4m}z^2}{(w-q^{2m}z)^2} + \frac{q^{4n}\zeta^2}{(w-q^{2n}\zeta)^2} - \frac{2q^{2m+2n}z\zeta}{(w-q^{2m}z)(w-q^{2n}\zeta)}\right\}.$$

With

$$\frac{1}{(w-q^{2m}z)(w-q^{2n}\zeta)} = \frac{1}{q^{2m}z-q^{2n}\zeta}\left\{\frac{1}{w-q^{2m}z} - \frac{1}{w-q^{2n}\zeta}\right\},$$

and introduction of the factor q^{2m+2n}, this becomes, after simple transformations of the exponents of q:

$$\begin{aligned} w^3\cdot\frac{q^{2m}}{(w-q^{2m}z)^2}\cdot\frac{q^{2n}}{(w-q^{2n}\zeta)^2} &= z^2\cdot\frac{q^{2n-2m}}{(z-q^{2n-2m}\zeta)^2}\cdot\frac{q^{4m}}{(w-q^{2m}z)^2} \\ &+ \zeta^2\cdot\frac{q^{2n-2m}}{(z-q^{2n-2m}\zeta)^2}\cdot\frac{q^{4n}}{(w-q^{2n}\zeta)^2} \\ &- z\zeta\cdot\frac{2q^{4n-4m}}{(z-q^{2n-2m}\zeta)^3}\cdot\frac{q^{2m}}{w-q^{2m}z} \\ &- z\zeta\cdot\frac{1}{(z-q^{2n-2m}\zeta)^3}\cdot\frac{q^{2n}}{w-q^{2n}\zeta}. \end{aligned}$$

Taking the derivative with respect to w:

$$\begin{aligned} \frac{\partial}{\partial w}\left\{w^2\cdot\frac{q^{2m}}{(w-q^{2m}z)^2}\cdot\frac{q^{2n}}{(w-q^{2n}\zeta)^2}\right\} &= -z^2\cdot\frac{q^{2n-2m}}{(z-q^{2n-2m}\zeta)^2}\cdot\frac{2q^{4m}}{(w-q^{2m}z)^3} \\ &- \zeta^2\cdot\frac{q^{2n-2m}}{(z-q^{2n-2m}\zeta)^2}\cdot\frac{2q^{4n}}{(w-q^{2n}\zeta)^3} \\ &+ z\zeta\cdot\frac{2q^{4n-4m}}{(z-q^{2n-2m}\zeta)^3}\cdot\frac{q^{2m}}{(w-q^{2m}z)^2} \\ &- z\zeta\cdot\frac{2q^{2n-2m}}{(z-q^{2n-2m}\zeta)^3}\cdot\frac{q^{2n}}{(w-q^{2n}\zeta)^2}. \end{aligned}$$

In this formula we allow the indices m, n to vary independently from $-\infty$ to $+\infty$ and sum; evidently the pairs of indices $n-m$, m and $n-m$, n then also vary independently from $-\infty$

to $+\infty$, and we may therefore replace them in the summation simply by m, n, writing

$$(8.1)\quad \frac{\partial}{\partial w}\left\{w^2 \cdot \sum_m \frac{q^{2m}}{(w-q^{2m}z)^2} \cdot \sum_n \frac{q^{2n}}{(w-q^{2n}\zeta)^2}\right\}$$
$$= -z^2 \cdot \sum_m \frac{q^{2m}}{(z-q^{2m}\zeta)^2} \cdot \sum_n \frac{2q^{4n}}{(w-q^{2n}z)^3}$$
$$-\zeta^2 \cdot \sum_m \frac{q^{2m}}{(z-q^{2m}\zeta)^2} \cdot \sum_n \frac{2q^{4n}}{(w-q^{2n}\zeta)^3}$$
$$+z\zeta \cdot \sum_m \frac{2q^{4m}}{(z-q^{2m}\zeta)^3} \cdot \sum_n \frac{q^{2n}}{(w-q^{2n}z)^2}$$
$$-z\zeta \cdot \sum_m \frac{2q^{2m}}{(z-q^{2m}\zeta)^3} \cdot \sum_n \frac{q^{2n}}{(w-q^{2n}\zeta)^2}.$$

Now, if for the moment we separate the logarithmic term from $P(z, \zeta)$ (formula 4.6), writing

$$P(z, \zeta) = -\frac{1}{2\log q} \cdot \frac{1}{z\zeta} + Q(z, \zeta),$$
$$Q(z, \zeta) = \sum_m \frac{q^{2m}}{(z-q^{2m}\zeta)^2},$$

then the terms of (8.1) are simply Q and its derivatives:

$$Q_z(z, \zeta) = \sum_m \frac{-2q^{2m}}{(z-q^{2m}\zeta)^3},$$
$$Q_\zeta(z, \zeta) = \sum_m \frac{2q^{4m}}{(z-q^{2m}\zeta)^3};$$

accordingly (8.1) may be written:

$$(8.1')\quad \frac{\partial}{\partial w}\{w^2 Q(z, w)Q(\zeta, w)\}$$
$$= -z^2 Q(z, \zeta)Q_z(z, w) - \zeta^2 Q(z, \zeta)Q_\zeta(\zeta, w)$$
$$+ z\zeta Q(z, w)Q_\zeta(z, \zeta) + z\zeta Q(\zeta, w)Q_z(z, \zeta).$$

If then we substitute back in this

$$Q(z, \zeta) = P(z, \zeta) + \frac{1}{2\log q} \cdot \frac{1}{z\zeta},$$

it turns out that, curiously enough, *we get the same identity in* P:

$$(8.2)\quad \frac{\partial}{\partial w}\{w^2P(z,w)P(\zeta,w)\} = -z^2P(z,\zeta)P_z(z,w) - \zeta^2P(z,\zeta)P_\zeta(\zeta,w) + z\zeta P(z,w)P_\zeta(z,\zeta) + z\zeta P(\zeta,w)P_z(z,\zeta).$$

The principle point in the calculation is the cancellation of

$$zwP(z,w) + \zeta wP(\zeta,w) + zw^2P_w(z,w) + \zeta w^2P_w(\zeta,w)$$

on the left, with

$$z\zeta P(z,\zeta) - z^2wP_z(z,w) + z\zeta P(z,\zeta) - \zeta^2wP_\zeta(\zeta,w) - zwP(z,w) + z\zeta^2P_\zeta(z,\zeta) - \zeta wP(\zeta,w) + z^2\zeta P_z(z,\zeta)$$

on the right, in virtue of the Euler equation for the homogeneous function P of degree -2:

$$zP_z(z,\zeta) + \zeta P_\zeta(z,\zeta) + 2P(z,\zeta) = 0,$$

with the analogous equations in z, w and ζ, w.

Next, we integrate both sides of (8.2) as to w, and get

$$w^2P(z,w)P(\zeta,w) = -z^2P(z,\zeta)Z_z(z,w) - \zeta^2P(z,\zeta)Z_\zeta(\zeta,w) + z\zeta Z(z,w)P_\zeta(z,\zeta) + z\zeta Z(\zeta,w)P_z(z,\zeta) + \phi(z,\zeta),$$

the last term being the constant of integration. It can be evaluated by identifying the constant terms in the Laurent expansion as to w of the two members of the equation; with the use of the Laurent expansion of $P(z,\zeta)$:

$$P(z,\zeta) = \frac{1}{z\zeta}\sum_{m=-\infty}^{\infty}\frac{m}{1-q^{2m}}\cdot\frac{z^m}{\zeta^m}\left(\text{for } m=0,\ \frac{m}{1-q^{2m}} = -\frac{1}{2\log q}\right),$$

we thus find

$$\phi(z,\zeta) = \frac{1}{2}q\frac{\partial}{\partial q}P(z,\zeta),$$

and have the desired identity (7.1).

9. Interlacing Contours

Let $\bar{g}_1$ minimize the functional $A(g)$ for the contour Γ_1 and $\bar{g}_2$ for the contour Γ_2; then $\bar{g}_1$ and $\bar{g}_2$ determine respectively a minimal surface M_1 bounded by Γ_1 and a minimal surface M_2 bounded by Γ_2. Without alteration of M_1 or M_2, $\bar{g}_1$ and $\bar{g}_2$ are each subject to arbitrary linear transformation $z' = \dfrac{az + b}{cz + d}$ of the circles C_1 and C_2 into themselves respectively. We can adjust the three constants involved so that any chosen point and (sensed) direction on M_1 correspond by the equations of M_1, $x_i = \Re F_{1i}(w)$, to the origin $w = 0$ and the direction of the positive real axis; and similarly for M_2.

By the expansion [15]

$$F_{1i}(w) = \frac{a_{1i0}}{2} + (a_{1i1} - ib_{1i1})w + (a_{1i2} - ib_{1i2})w^2 + \cdots,$$

we see that the point $w = 0$ and the unit vector in the direction of the positive real axis go over, by the equations of M_1, respectively into the point P of coordinates

$$P: \qquad \frac{1}{2\pi}\int_0^{2\pi} \bar{g}_{1i}(\theta)\,d\theta$$

and the vector a_{1i1}; the unit vector along the imaginary axis is converted into the vector b_{1i1} perpendicular to a_{1i1}; both these vectors have the same length μ_1:

$$\mu_1 = \sqrt{\sum_i (a_{1i1})^2} = \sqrt{\sum_i (b_{1i1})^2} = \sqrt{\tfrac{1}{2}\sum_i |F_{1i}'(0)|^2},$$

representing the linear stretching factor of the conformal transformation from the circular disc to the minimal surface. We suppose $\mu_1 > 0$. Under this condition the point P is called regular; otherwise it is singular. The singular points are necessarily isolated; there are at most a finite number of them in any region contained with its boundary in the interior of M_1.

[15] Refer back to § 5 for the notation of Fourier coefficients.

Suppose now that *the minimal surfaces M_1 and M_2 have a point P in common that is regular for both of them*:

$$\sum_{i=1}^{n} |F_{1i}'(w)|^2 > 0, \qquad \sum_{i=1}^{n} |F_{2i}'(w)|^2 > 0.$$

We can then dispose of two of the constants involved in $\bar{g}_1$, $\bar{g}_2$ respectively to make P correspond to $w = 0$ on both M_1 and M_2; thus

$$\frac{1}{2\pi}\int_0^{2\pi} \bar{g}_{1i}(\theta)d\theta = \frac{1}{2\pi}\int_0^{2\pi} \bar{g}_{2i}(\theta)d\theta. \tag{9.1}$$

We next make the supposition that *the tangent planes at P to M_1 and M_2 are not absolutely perpendicular.* Absolutely perpendicular means that every direction in the one is perpendicular to every direction in the other. This supposition may be omitted if we restrict the problem to three-dimensional space, being then satisfied of itself.

Let then two directions making with one another an angle $\alpha < \frac{\pi}{2}$ be selected respectively in the tangent planes at P; by making use of the third arbitrary constant in $\bar{g}_1$, $\bar{g}_2$ we can arrange that the vectors a_{1i1}, a_{2i1} take these prescribed directions; then their scalar product

$$\sum_i a_{1i1}a_{2i1} = \mu_1\mu_2 \cos\alpha > 0. \tag{9.2}$$

With a possible reflection in the real axis, we can then arrange that the vectors b_{1i1}, b_{2i1} make with each other an angle $\beta \leqq \frac{\pi}{2}$ so that

$$\sum_i b_{1i1}b_{2i1} = \mu_1\mu_2 \cos\beta \geqq 0; \tag{9.2'}$$

hence with (9.2):

$$\sum_i (a_{1i1}a_{2i1} + b_{1i1}b_{2i1}) > 0. \tag{9.3}$$

Thus the coefficient of q in the expansion (5.6) is negative, strictly < 0. But also, by (9.1) $\alpha = 0$ (refer to 5.2), so that

the term containing the reciprocal of $-\log q$ disappears and the term in q becomes dominant for all sufficiently small q; therefore for q small enough:

$$A(\bar{g}_1, \bar{g}_2; q) < A(\bar{g}_1) + A(\bar{g}_2) = m(\Gamma_1) + m(\Gamma_2).$$

Consequently,

$$m(\Gamma_1, \Gamma_2) < m(\Gamma_1) + m(\Gamma_2),$$

our basic sufficient condition, and we have the following theorem:

THEOREM II. *In euclidean space of n dimensions, let Γ_1, Γ_2 be two Jordan curves which do not intersect one another and for which $m(\Gamma_1, \Gamma_2)$ is finite.*

If the minimal surfaces M_1 and M_2 determined by Γ_1 and Γ_2 taken separately have in common a point P that is regular for both of them and where the tangent planes to M_1 and M_2 are not absolutely perpendicular, then there exists a doubly connected minimal surface M_{12} bounded by Γ_1 and Γ_2.

The area of M_{12} is less than the sum of the areas of M_1 and M_2.

As has been remarked, in the most important case $n = 3$ the condition that the tangent plánes be not absolutely perpendicular is satisfied automatically, and may be omitted from the statement of the theorem.

An interesting and immediate corollary of this theorem refers to the case when the two contours interlace. It is evident, as a matter of topology, that the surfaces M_1 and M_2 then have in common a point P, non-singular because it may be taken anywhere on the (analytic) curve of intersection of M_1 and M_2, whereas the singular points of these surfaces are isolated. Hence:

THEOREM III. *In three-dimensional euclidean space, let Γ_1 and Γ_2 be Jordan curves with finite $m(\Gamma_1, \Gamma_2)$ which interlace.*

Then Γ_1 and Γ_2 are the boundaries of a doubly-connected minimal surface.

The area of this surface is less than the sum of the areas of the simply-connected minimal surfaces determined by Γ_1 and Γ_2 separately.

10. Conformal Mapping

The conformal mapping of a doubly-connected plane region bounded by two Jordan curves on a circular ring is nothing but the special case $n = 2$ of the preceding theory.

That a Jordan curve in the plane has finite $m(\Gamma)$ is known from (I), § 24; in fact, $m(\Gamma)$ is equal to the inner area of the region enclosed by Γ (upper bound of the areas of polygons which, with their boundaries, lie in the interior of the region). Hence $m(\Gamma_1, \Gamma_2) \leqq m(\Gamma_1) + m(\Gamma_2)$ is finite.

The minimal surfaces M_1 and M_2 of Theorem II are here simply the plane regions enclosed by Γ_1 and by Γ_2, so that all the hypotheses of that theorem are satisfied if we take for P any point interior to the region enclosed by the inner contour Γ_2.

Therefore we can conclude the existence of two harmonic functions

$$x_1 = \Re F_1(w), \qquad x_2 = \Re F_2(w), \tag{10.1}$$

where F_1 and F_2 are homomorphic in the interior of a certain circular ring C_1, C_2 and obey there identically the relation

$$F_1'^2(w) + F_2'^2(w) = 0. \tag{10.2}$$

Moreover, the transformation (10.1), from w to (x_1, x_2) remains continuous on the boundaries C_1, C_2 and converts them in a topological way into Γ_1 and Γ_2.

The condition (10.2) implies that the functions (10.1) are conjugate harmonic, so that

$$x_1 + ix_2 = F(w), \tag{10.3}$$

monogenic and regular in the circular ring. From the fact that $F(w)$ converts the bounding circles in a one-one continuous way into Γ_1, Γ_2, it follows by elementary facts of complex function theory (I, § 23) that $F(w)$ establishes a one-one and conformal map (no singularities, $F'(w) \neq 0$) between the circular ring and the ring between Γ_1 and Γ_2.

Thus we have a result combining the classic conformal mapping theorem of Schottky relating to the interiors of doubly-connected regions with the theorem of Osgood relating to the behavior of the conformal map on the boundaries.

We summarize as follows:

THEOREM IV. *Let Γ_1, Γ_2 denote any two Jordan curves in the plane which enclose between them a region R. In the equations*

$$Z = g_1(z), \qquad Z = g_2(z),$$

where Z and z denote complex variables, let Z describe Γ_1, Γ_2 respectively when z describes the two concentric circles C_1, C_2: $|z| = 1$, $|z| = q$ with $0 < q < 1$.

The range of values of the functional

$$A(g_1, g_2; q) = \frac{1}{4\pi} \sum_{\alpha\beta} \int_{C_\alpha} \int_{C_\beta} |g_\alpha(z) - g_\beta(\zeta)|^2 P(z, \zeta; q) dz d\zeta,$$

when all parametric representations g_1, g_2 of Γ_1, Γ_2 and all values of q are considered, will consist exactly of all positive real numbers $\geqq$ the inner area of the region R. This minimum value will be attained for a certain (essentially uniquely determined) parametric representation

$$Z = g_1^*(z), \qquad Z = g_2^*(z)$$

together with a unique value q^* of q.

Then the integral formula of Cauchy:

$$W = \frac{1}{2\pi i} \int_{C_1} \frac{g_1^*(z) dz}{z - w} + \frac{1}{2\pi i} \int_{C_1} \frac{g_2^*(z) dz}{z - w}$$

defines a conformal transformation $w \to W$ of the circular ring between C_1 and C_2 into the region R between Γ_1 and Γ_2; and this conformal transformation attaches continuously to the topological correspondence g_1^*, g_2^* between the boundaries.

11. Arbitrary Jordan Contours

If we separate from the expansion (4.6) of $P(z, \zeta)$ the term $\frac{1}{(z-\zeta)^2}$, writing

$$P(z, \zeta) = \frac{1}{(z-\zeta)^2} + R(z, \zeta), \tag{11.1}$$

$$R(z, \zeta) = -\frac{1}{2 \log q} \cdot \frac{1}{z\zeta} + \sum_{m=-\infty}^{\infty}{}' \frac{q^{2m}}{(z - q^{2m}\zeta)^2}, \tag{11.1'}$$

the accent attached to the summation sign indicating that m is not to take the value zero; then, for q bounded away from 1, the function $R(z, \zeta)$ stays finite as z, ζ vary independently over the circles C_1, C_2, and we are led to write

$$A(g_1, g_2; q) = A(g_1) + A(g_2) - B(g_1, g_2; q), \tag{11.2}$$

or

$$B(g_1, g_2; q) = A(g_1) + A(g_2) - A(g_1, g_2; q), \tag{11.2'}$$

with

$$\begin{aligned} B(g_1, g_2; q) = &-\frac{1}{4\pi} \sum_{\alpha\beta} \int_{C_\alpha} \int_{C_\beta} \sum_{i=1}^{n} [g_{\alpha i}(z) - g_{\beta i}(\zeta)]^2 \\ &\times R(z, \zeta; q) dz d\zeta \\ &- \frac{1}{2\pi} \int_{C_1} \int_{C_1} \sum_{i=1}^{n} [g_{1i}(z) - g_{2i}(\zeta)]^2 \frac{dz d\zeta}{(z-\zeta)^2}. \end{aligned} \tag{11.2''}$$

B may be called the finite part of the functional A, for such is its character, since its integrand stays bounded provided q is bounded away from 1; only the terms $A(g_1)$, $A(g_2)$ can cause the value of $A(g_1, g_2; q)$ to be infinite. B is moreover a continuous functional, since the boundedness of its integrand is uniform provided the contours Γ_1, Γ_2 remain in a fixed finite region of space.

By these remarks we see that if either $A(g_1)$ or $A(g_2)$ is identically $+\infty$, so also is $A(g_1, g_2; q)$, and conversely; in other words, if either $m(\Gamma_1)$ or $m(\Gamma_2) = +\infty$, then $m(\Gamma_1, \Gamma_2) = +\infty$, and conversely.

The problem now is to extend the preceding existence theory to the case $m(\Gamma_1, \Gamma_2) = +\infty$. To this end we introduce first the functional of two contours

$$e(\Gamma_1, \Gamma_2) = m(\Gamma_1) + m(\Gamma_2) - m(\Gamma_1, \Gamma_2), \tag{11.3}$$

defined by this formula for all pairs of contours with finite $m(\Gamma_1, \Gamma_2)$; we have seen, relation (1.4), that always $e(\Gamma_1, \Gamma_2) \geqq 0$. Based on this, we then define

$$\bar{e}(\Gamma_1, \Gamma_2) = \limsup e(\Gamma_1', \Gamma_2') \tag{11.4}$$

when Γ_1', Γ_2', contours with finite $m(\Gamma_1', \Gamma_2')$, approach Γ_1, Γ_2 in all the possible ways.[16] The functional $\bar{e}(\Gamma_1, \Gamma_2)$ is thus well-defined for every pair of contours, irrespective of whether $m(\Gamma_1, \Gamma_2)$ is finite or infinite; and always $\bar{e}(\Gamma_1, \Gamma_2) \geqq 0$.

We can now state

THEOREM V. *Any two Jordan curves Γ_1, Γ_2, not intersecting one another, for which*

$$\bar{e}(\Gamma_1, \Gamma_2) > 0$$

are the boundaries of a doubly-connected minimal surface.

As a preliminary to the proof, we develop certain inequalities (11.5, 11.6) for the case of finite $m(\Gamma_1, \Gamma_2)$. Consider the minimizing representation g_1^*, g_2^*; q^* of $A(g_1, g_2; q)$ and also $\bar{g}_1$, $\bar{g}_2$ which respectively minimize $A(g)$ for Γ_1 and for Γ_2. By (11.2′),

$$B(g_1^*, g_2^*; q^*) = A(g_1^*) + A(g_2^*) - A(g_1^*, g_2^*; q^*),$$
$$B(\bar{g}_1, \bar{g}_2; q) = A(\bar{g}_1) + A(\bar{g}_2) - A(\bar{g}_1, \bar{g}_2; q).$$

[16] If, for instance, Γ_1, Γ_2 are respectively the outer and inner boundaries of a ring-shaped plane region, and Γ_2 (the curve, or thread, itself) has positive area a, so that the inner and outer areas I, O of the region interior to Γ_2 differ by a; then, as is easy to see,

$$\bar{e}(\Gamma_1, \Gamma_2) = \limsup e(\Gamma_1', \Gamma_2') = 2O,$$
$$\underline{e}(\Gamma_1, \Gamma_2) = \liminf e(\Gamma_1', \Gamma_2') = 2I,$$
$$e(\Gamma_1, \Gamma_2) = I + O.$$

The approaching contours Γ_1', Γ_2' are supposed to remain in the plane.

In the first of these formulas, the negative term has its minimum value, while the positive terms are $\geqq$ their minimum values; hence

$$(11.5)\quad B(g_1^*, g_2^*; q^*) \geqq m(\Gamma_1) + m(\Gamma_2) - m(\Gamma_1, \Gamma_2) = e(\Gamma_1, \Gamma_2).$$

In the second formula, the positive terms have their minimum values, while the negative term is $\geqq$ its minimum: hence, for all q,

$$(11.6)\quad B(\bar{g}_1, \bar{g}_2; q) \leqq m(\Gamma_1) + m(\Gamma_2) - m(\Gamma_1, \Gamma_2) = e(\Gamma_1, \Gamma_2).$$

Let Γ_1', Γ_2' be a variable pair of contours with finite $m(\Gamma_1', \Gamma_2')$ tending to Γ_1, Γ_2 and such that $m(\Gamma_1', \Gamma_2')$ tends to its superior limit:

$$(11.7)\qquad \lim e(\Gamma_1', \Gamma_2') = \bar{e}(\Gamma_1, \Gamma_2) > 0.$$

Then at every stage in its variation Γ_1', Γ_2' has a minimizing representation $'g_1^*, 'g_2^*; 'q^*$. By the compactness of the totality of parametric representations of a given contour, as used in (I), § 19, we can select a sub-sequence of Γ_1', Γ_2' so that $'g_1^*$ and $'g_2^*$ tend respectively to proper, improper or degenerate representations g_1° and g_2° of Γ_1 and Γ_2, while $'q^*$ tends to a limit $q_0 \geqq 0$ and $\leqq 1$.

We must first exclude the possibility of the extreme values 0 and 1 for q_0.

To obviate the possibility $q_0 = 0$, we observe that, since $'g_1^*$, $'g_2^*$ remain uniformly bounded, then if $'q^* \to 0$, we have by (11.2″), (11.1′),

$$B('g_1^*, 'g_2^*; 'q^*) \to 0$$

(the last term in (11.2″) approaching to zero because the circle C_2 of radius $'q^*$ shrinks to a point). According to this, it would follow from (11.5) that the non-negative quantity $e(\Gamma_1', \Gamma_2')$ approaches to zero; i.e. by (11.7),

$$\bar{e}(\Gamma_1, \Gamma_2) = 0,$$

contrary to hypothesis.

To eliminate $q_0 = 1$, we remark that $'q^*$ is a root of the equation

$$'\phi(q) = \frac{\partial A('g_1^*, 'g_2^*, q)}{\partial q} = 0,$$

and by (5.8) we have for $q \to 1$:

$$'\phi(q) \sim \frac{3'\alpha^2 + \frac{1}{2}\int_0^{2\pi} \sum_{i=1}^{n} ['g_i^*(\theta) - 'g_i^*(\theta)]^2 d\theta}{(1-q)^2};$$

moreover, this asymptotic expression applies uniformly to all the contour-pairs Γ_1', Γ_2' since these remain in a finite region of space. Let $d > 0$ denote the distance between Γ_1 and Γ_2, and d_1 any positive value $< d$; then from a certain stage on in the variation of Γ_1', Γ_2' we have

$$'\phi(q) > \frac{\pi d_1^2}{(1-q)^2},$$

which makes it impossible for $'\phi(q)$ to have a root tending to 1.

Thus q_0 is > 0, < 1, and gives a proper circular ring C_1, C_2.

The harmonic functions $x_i = \Re F_i(w)$ in this circular ring with the boundary values g_1°, g_2° must obey the condition for a minimal surface:

$$\sum_{i=1}^{n} F_i'^2(w) = 0,$$

since they are the limits of functions always obeying this condition.

That the representations g_1°, g_2° cannot be improper results by the same arguments as in (I), §§ 18, 20. First, g_1°, g_2° cannot be improper of types 2° or 3° (see § 6), for, as has been said, under the condition $\sum_{i=1}^{n} F_i'^2(w) = 0$, the vector function $\Re F_i(w)$ cannot reduce to a constant on an arc of C_1 or C_2 without reducing to a constant identically; now this cannot be, because the contours Γ_1, Γ_2 are supposed to have no point in common.

The possibility of g_1° or g_2° being improper of type 1° can be ruled out, as in (I), § 20, by an inversion of the circular ring with respect to a circle orthogonal to the one which carries the supposed point P of discontinuity, for instance C_1. If the circle of inversion, always remaining orthogonal to C_1, shrinks towards the point P, then C_2 shrinks towards the same point, and its contribution to the value of the harmonic function $\Re F_i(w)$ becomes less and less indefinitely, so that we have essentially the same situation as in the one-contour case.

In sum, we have proved Theorem V.

As a corollary to Theorem V, we conclude by extending Theorems II and III to arbitrary Jordan contours.

It is first to be remarked that from the inequality (11.6), established for finite $m(\Gamma_1, \Gamma_2)$, we can easily deduce

$$B(\bar{g}_1, \bar{g}_2; q) \leqq \bar{e}(\Gamma_1, \Gamma_2) \tag{11.8}$$

for all q, applicable also in the case of infinite $m(\Gamma_1, \Gamma_2)$, provided $\bar{g}_1$, $\bar{g}_2$ are interpreted as representations of Γ_1, Γ_2 which give the minimal surfaces M_1 and M_2 determined by Γ_1 and Γ_2 separately; the existence of $\bar{g}_1$, $\bar{g}_2$ was proved in (I), Part II. For we can arrange a sequence of contours Γ_1', Γ_2' with finite $m(\Gamma_1', \Gamma_2')$, tending to Γ_1, Γ_2 and such that

$$\lim e(\Gamma_1', \Gamma_2') = \bar{e}(\Gamma_1, \Gamma_2),$$

while the separate minimizing representations $'\bar{g}_1$, $'\bar{g}_2$ of Γ_1', Γ_2' tend to $\bar{g}_1$, $\bar{g}_2$. Then because the functional B is continuous,

$$\lim B('\bar{g}_1, '\bar{g}_2; q) = B(\bar{g}_1, \bar{g}_2; q)$$

for any fixed q. But since always during the limit process, by (11.6),

$$B('\bar{g}_1, '\bar{g}_2; q) \leqq e(\Gamma_1', \Gamma_2'),$$

it follows that

$$B(\bar{g}_1, \bar{g}_2; q) \leqq \bar{e}(\Gamma_1, \Gamma_2) \tag{11.8}$$

for all q, which was to be proved.

Now if the minimal surfaces M_1, M_2 have a regular point in common, where the tangent planes are not absolutely perpendicular, then, exactly as in § 9, with simply our new notation B, i.e.:

$$B(g_1, g_2; q) = \frac{4\pi\alpha^2}{\log q} + 2\pi q(a_{1,1}a_{2,1} + b_{1,1}b_{2,1}) + \cdots$$

(being (5.6) rewritten in accordance with (11.2)), we have

$$B(\bar{g}_1, \bar{g}_2; q) > 0$$

for all sufficiently small q. Hence by (11.8), $\bar{e}(\Gamma_1, \Gamma_2) > 0$, the sufficient condition of Theorem V, and we therefore have:

THEOREM VI. *In euclidean space of n dimensions, let Γ_1, Γ_2 be two Jordan curves which do not intersect one another.*

If the minimal surfaces M_1 and M_2 determined by Γ_1 and Γ_2 taken separately have in common a point P that is regular for both of them and where the tangent planes to M_1 and M_2 are not absolutely perpendicular, then there exists a doubly-connected minimal surface M_{12} bounded by Γ_1 and Γ_2.

In the most important case $n = 3$ the condition that the tangent planes at P be not absolutely perpendicular is satisfied of itself and may be omitted from the statement of the theorem.

As an immediate corollary to Theorem VI there follows, as in § 9, the following theorem with respect to interlacing contours.

THEOREM VII. *In three-dimensional euclidean space, let Γ_1 and Γ_2 be any two Jordan curves which interlace.*

Then Γ_1, Γ_2 are the boundaries of a doubly-connected minimal surface.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY,
CAMBRIDGE, MASS.,
Nov.–Dec., 1931.

Bibliography

Vocabulary